

34 InSerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50
|||||
153 AATCTACCAAGTACAAGGCAGACAAACCTATCTCAACTGTGGCTGAG 202
51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
::: |||||
203 AAGCCCAAGAATATCAAGAAAAACAGATATAAGGATATTTGGCCCTATGA 252
67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSert 84
|||||
253 TTATAGCCGGGTAGAAGTATCCCTGATAACCTCTGATGAGGATCCAGCT 302
84 yrIleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
|||||
303 ACATCAATGCCAACTTCATTAAAGGAGTTTATGGACCCRAGGCTTATATT 352
101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetIl 117
|||||
353 GCCACCACAGGCTCTTATCTACAAACCTCCTGGACTTCTGGAGGATGAT 402
117 eTTPGluTyrSerValLeuIleIleValMetAlaCysMetGluTyrGluM 134
|||||
403 TTGGGAATATAGTGCTCTTATCATTTGTTATGGCATGCATGGAGTATGAAA 452
134 etGlyLysLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
|||||
453 TGGGAAGAAAAAGTGTGACGGCTACTGGCTGAGCCAGCAGAGATGCAG 502
151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
|||||
503 CTGGAATTGGCCCTTCTCTGTATCTCTGTGAAGCTGAAAAAGGAAATC 552
167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
|||||
553 TGATATATATAATCAGGACTCTAAAGTTAAGTTCAATAGTGAACATCGAA 602
184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSer 200
|||||
603 CTATCTACCAAGTTTCATTACAAGATTGGCCAGACCATGATGTACCTTCA 652
201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnGl 217
|||||
653 TCTATAGACCCCTATTCTGTAGCTCATCTGGGATGTACGTTGTACCAGA 702
217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
|||||
703 GGATGACAGTGTCCCATATGCATTTCACTGCAGTGTGGCTGTGGGAAGA 752
234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
|||||
753 CTGGTGTATTGTGCTATT...GATTATACATGGATGTGTCTAAAGAT 799
251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267
|||||
800 GGGATAATTCTCTGAGAACTTCAGTGTGTTTCAGTTGATCCGGGAAATGCG 849
267 qThrGlnArgProSerLeuValGlnThrGlnGluInTyrGluLeuValT 284
|||||
850 GACACAGAGGCCCTTCATTAGTCTAAACGCAGAGAACATATGAATCTGGTCT 899
284 yrAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
|||||
900 ACAATGCTGTATTAGAACTATTAAAGACACAGATGGATGTTATCAGAGAT 949
301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
|||||
950 AAACATTCTGGAAACAGAGATCAAGCAAAAGCATTTGTATTCCTCTGAGAAAA 999
317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
|||||
1000 TCACACTCTCCAGACAGACTCTTATTCTCTTAATTTACCAAAAAGTACCA 1049
334 hrLysAlaAlaLysMetMetAsnGlnArgThrLysMetGluIleLys 350
|||||

1050 CAAAACGACGCAAAAATGATGAACCAACAAAGGACAAAATGGAAATCAAA 1099
351 GluSerSerSerPheAspPheArgThrSerGluIleSerAlaLysGluGl 367
|||||
1100 GAATCTCTCTCTCTTACTTACTTAGGACTTCTGAAATTAAGTGCAAAAAGA 1149
367 uLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGluL 384
|||||
1150 GCTAGTGTTCACCCCTGCTAATCAACAGCCTCTCTTTGACTTTCTGGAGC 1199
384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnThr 400
|||||
1200 TAAATACAGTCTTTGACAAAAATGCTGACACAACCATGAATGGCAGACA 1249
401 LysAlaPheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAs 417
|||||
1250 AAGGCATTTTCCAATAGTTGGGGAGCCTCTTTCAGAAGCATCAAAAGTTTGA 1299
417 pLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsnA 434
|||||
1300 TTTGGGCTCTCTTTTGTGTTGAGGATGTTCTTAATTTCTAAACCTGTAATG 1349
434 laAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSer 450
|||||
1350 CACGACGAGAGATATTTAATTCAAGGTGCAATAACACGGACCAAAATCA 1399
451 ThrProPheGluLeuIleGlnArgGluThrLysGluValAspSerLy 467
|||||
1400 ACTCTCTTTTGAATTGATACAGCAGAGAGAAACCAAGAGGTGACAGCAA 1449
467 sGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheValG 484
|||||
1450 GGAAGACTTTTCTATTGGAATCTCAACCATCATGATCTTGTTTGTAG 1499
484 luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
|||||
1500 AGATGACGAGCTCAAAAAGTAATGCATGTTTCTTCAGCAGAACTGAATAT 1549
501 SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnValLy 517
|||||
1550 TCACCTGCCATATGACTCTAAACACCAAAATACGTAATGCCCTCTAATGTAAA 1599
517 sHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValG 534
|||||
1600 GCACCATGACTCTAGTGCTCTTGGTGTATATTCTTACATACCTTTAGTGG 1649
534 luAsnProTyrPheSerSerTrpProProSerGlyThrSerSerLysMet 550
|||||
1650 AAAATCCTTATTTTTCATCATGCTCCCAAGTGTACCAGTTCTAAGATG 1699
551 SerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerLe 567
|||||
1700 TCTCTTGATTTTACCTGAGAGCAAGATGGAACTGTTTTTCTCTCTCTCT 1749
567 uLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSerS 584
|||||
1750 GTTGGCCACATCTCTACATCCCTCTCTCTCTTATTACAAATTCACATGATT 1799
584 erLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnGlu 600
|||||
1800 CTTTATCACTGAATTCTCCAACCAATATTTCTCTCACTATTATGAACCCAGGAG 1849
601 SerAlaValLeuAlaThrAlaProArgIleAspAspGluIleProProPr 617
|||||
1850 TCAGCTGTACTAGCAACTGCTCCAAGGATAGATGATAAATCCCCCTCC 1899
617 oLeuProValArgThrProGluSerPheIleValValGluGluAlaGlyG 634
|||||
1900 ACTTCTGTACGACACACTGAATCAITTTATTTGGTTGAGGAAGCTGGAG 1949
634 luPheSerProAsnValProLysSerLeuSerSerAlaValLysValLys 650
|||||

[illegible]

seq_documentation block:

; Patent No. 6004791

APPLICANT: Aoki,

ATTORNEY: OFFICE
; TITLE OF INVENTION

FILE OF INVENTION	NUMBER OF SEQUENCE
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

; ADDRESS: 1 von
; CORRESPONDENCE ADD

STREET: 633 West
CITY: Salt Lake City, UT 84143

CITY: Los Angeles

COUNTRY: U.S.A.

; COMPUTER READABLE

;
MEDIUM TYPE: ST

; OPERATING SYSTEM

1. CURRENT APPLICATION

FILING DATE: 0C

7. MODIFICATION:
: PRIOR APPLICATION

APPLCATION NUMBER: _____
FILING DATE: _____

APPLICANT NAME: _____
FILING DATE: _____

ATTORNEY/AGENT INFO
NAME: Warburg

:	REGISTRATION NUM
:	REFERENCE / DOCKET

TELECOMMUNICATIONS;

TELEFAX: (213) 573-2510

; INFORMATION FOR SEQ

; LENGTH: 2226 ba

; STRANDEDNESS: S

; MOLECULE TYPE: CD

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

argument_scores:

Percent Similarity

alignment block.

US-09-600-358A-4

```

881 ACCACAGTGGTCACTATTCTCCGCACACTCTCCAGAACAAACAGTCCC 930
297 .....ValileArgAspLysH1 302
931 CTCTACCAGAACCTCAAGGAGAACCGGCTCCATCTGCAAGGACTCCTC 980
302 sSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAsnH1st 319
111 .....CTGCCTGCCACATCCCGCC 1021
981 GTCCCTCAGGACCTCCACGCC.....CTGCCTGCCACATCCCGCC 1021
319 hrLeuGlnAlaAspSerTyrSer.....ProAsnLeu 329
111 ..... 1022
1022 CACTGGGTGGCTTCTCAGGAGCATCTCGGTGCTGGGCCACCGACCTT 1071
330 ProLysSerThrThrLysAla 336
1072 CCCATGGCTGACACTTACGCT 1092

```

seq_name: /cqn2_6/ptodata/1/ina/6B_COMB.seq:US-08-821-278A-1

seq_documentation_block:

```

; Sequence 1, Application US/08821278A
; Patent No. 6238902
; GENERAL INFORMATION:
; APPLICANT: Cheng, Jill
; APPLICANT: Lasky, Laurence A.
; TITLE OF INVENTION: Protein Tyrosine Phosphatases
; FILE REFERENCE: P1010R1
; CURRENT APPLICATION NUMBER: US/08/821,278A
; CURRENT FILING DATE: 1997-03-20
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 1
; LENGTH: 1529
; TYPE: DNA
; ORGANISM: Mus musculus
US-08-821-278A-1

```

alignment_scores:

```

Quality: 736.50 Length: 327
Ratio: 3.069 Gaps: 6
Percent Similarity: 73.394 Percent Identity: 44.037

```

alignment_block:

```

US-09-600-358A-4 x US-08-821-278A-1 ..
Align seg 1/1 to: US-08-821-278A-1 from: 1 to: 1529
5 GluIleLeuGlnLysPheLeuAspGluAlaGlnSerLysLysIleThrLy 21
111 ..... 34
34 GACTTGGTGAGGAGCTTCTGGAGCAGCTGGAGCCCGGAGCTACCGGA 83
21 sGluGlu...PheAlaAsnGluPheLeuLysLeuLysArgGlnSerThrL 37
111 ..... 84
84 GGGGGCAATCTTCGTGCTGAGTTCAGCGACATTAAGGCCCGCTCAGTGG 133
37 ysTyrLysAlaAspLysThrThrProThrValAlaGluAsnAlaLys 53
111 ..... 134
134 CCTGGAAGTGTGAAGGTGTGTGTTCCACTAAAGCCGCGCAGTCGGCTGGG 183
54 AsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAspTyrSerAr 70
111 ..... 184
184 AACAGCAACAGAACCCGCTACAAAGATGTGGTAGCATATGATGAGACAAG 233
70 gValGluSerLeuIleThrSerAspGluAspSerTyrIleAsnA 87
111 ..... 234
234 AGTCATCCTTCTCCCTGCTCCAAAGAGGAGGACATGAAATATACATCAATG 283
87 laAsnPheIleLysGlyValTyrGlyProLysAlaTyrIleAlaThrGln 103
111 ..... 284
284 CCAACTTCATCCGGGGGATGAGTGAAGCCAGCCGCTACATGCGACGCAA 333
104 GlyProLeuSerThrThrLeuLeuAspPheTrpArgMetIleTrpGluTy 120

```

```

111 ..... 334
334 GGACCCCTGCCTCACACACTGTGTGGACTTCTGGCGCTGTGTGGAGTTF 383
120 rSerValLeuIleValMetAlaCysMetGluTyrGluMetGlyLysL 137
111 ..... 384
384 TGGGGTCAAGGTAATCCTGATGCCCTGTCAAGAGACAGAAAATGGACGGA 433
137 ysLysCysGluArgTyrTrpAlaGluProGlyGluMetGlnLeuGluPhe 153
111 ..... 434
434 GGAAGTGTGAACCTATTTGGGCCCGGGAGCAGGAG...CCTCTAAAGGCT 480
154 GlyProPheSerValSerCysGluAlaGluLysArg...LysSerAspTy 169
111 ..... 481
481 GGGCCTTCTGTCATCACCTCGCAAGAGGACAACTGAATGCAGACAT 530
169 rIleIleArgThrLeuLysValLysPheAsnSerGluThrArgThrIleT 186
111 ..... 531
531 CACTCTCAGGACCTCCAGGTTACATTTCCAGAAAGGAATTCCTGCTGTGC 580
186 yrGlnPheHisTyrLysAsnTrpProAspHisAspValProSerSerIle 202
111 ..... 581
581 ACCAACTACAGTATATGCTCTGCCAGACCCAGGGGTTCACGACGATCT 630
203 AspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnGluAspAs 219
111 ..... 631
631 GATCACATCTCTACCATGTTGGAGGAGGCCGCTGCCCAAGGGCTTGG 680
219 pSerValProIleCysIleHisCysSerAlaGlyCysGlyArgThrGlyv 236
111 ..... 681
681 ACCTGGACCCCTGTGTGTCACATGTCAGTGTGCTGCGGACGACACAGGTG 730
236 alIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAspGlyIle 252
111 ..... 731
731 TCCTGTGGCT...GTGACTATGTGAGGAGTGTGCTGCTGACCCACAGCA 777
253 IleProGluAsnPheSerValPheSerLeuIleArgGluMetArgThrGl 269
111 ..... 778
778 ATCCCTCCCAACTTCAGTCTCTTCCAAAGTGTCTCTGGAGATGCGGAAACA 827
269 nArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValTyrAsnA 286
111 ..... 828
828 GCGGCTGCGACGATGCGAGACAGAGGAGTACAGGTCTCTGTATACCACA 877
286 laValLeuGluLeuPheLysArgGlnMetAspValIleArgAspLysHis 302
111 ..... 878
878 CAGTGGCTCAGCTATCTCCGCGC.....ACTCTCCAGGACACACAGC 918
303 SerGlyThrGluSerGlnAlaLysHisCysIlePro.....GluLysAs 317
111 ..... 919
919 CCCCAATACCAAGAACCTCAAGGAGAACTGCGCTCCAATCTGCAAGGAAGC 968
317 nHisThrLeuGlnAlaAspSerTyrSerPro 327
111 ..... 969
969 TTTTCTCCCTCAGGACCTCTCTCAGCCCTGCGCT 999
seq_name: /cqn2_6/ptodata/1/ina/6A_COMB.seq:US-08-854-585-1
seq_documentation_block:
; Sequence 1, Application US/08854585
; Patent No. 6114140
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas K. and stman, Arne
; TITLE OF INVENTION: Density Enhanced Protein Tyrosine Phosphatase
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, Suite 6300
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```



```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,585
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/237,940
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27866/31954
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5117 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 350..4364
; US-08-854-585-1

alignment_scores:
    Quality: 508.00      Length: 295
    Ratio: 2.579         Gaps: 7
    Percent Similarity: 66.780      Percent Identity: 38.305

alignment_block:
US-09-600-358a-4 x US-08-854-585-1 ..
Align seg 1/1 to: US-08-854-585-1 from: 1 to: 5117

15  GlnSerLysLysIleThrLysGluLupheAlaAsnGluPheLeuLysLe 31
3398  AATCTAAGTTAATGACAGTGCGAAGATTTTGAGCGCTACTTCAAGAAGCA 3447
31  uLysArgGlnSerThr.....LysTyrLysAlaAspLys. 42
3448  GCAAGCTGACTCCACTGGTTCGCGAGAGGAATACGAAGATCTGAAGC 3497
43  .....ThrTyrProThrThrValAlaGluAsnAlaLysAsnIle 55
3498  TTGTTGGAATTAGTCAACCTAAATATGCAGCAGAACTGGCTGAGAATAGA 3547
56  LysLysAsnArgTyrLysAspIleLeuProTyrAspTyrSerArgValGl 72
3548  GGAAGAATCGCTATAATAATGTTCTGCCTATGATATTTCGCGTGTCAA 3597
72  uLeuSerLeuIleThrSerAspGluAspSerSerTyrIleAsnAlaAsp 89
3598  ACTTTCGGTCCAGACCCATTCAACGGAT...GACTACATCAATGCAACT 3644
89  heLysGlyValTyrGlyProLysAlaTyrIleAlaThrGlnGlyPro 105
3645  ACATCGCTGGCTACCACCTCCAAGAAGATTTTATGCCACAGAAGCACT 3694
106  LeuSerThrThrLeuLeuAspPheTrpArgMetIleTrpGluTyrSerVa 122
3695  TTACCGAACACATTTGAAAGATTTTGGCGTATGTTGGGAGAAAATGT 3744
122  IleuIleValMetAlaCysMetGluTyrGluMetGlyLysLysLysC 139
3745  ATATGCCATCATATGTGTGACTAATATGTTGACACAGGGAACCAAT 3794
139  ysGluArgTyrTrpAlaGluProGlyGluMetGlnLeuGluPheGlyPro 155
|||||:|||||  |||:|||||:  |||:|||||:

```

```

3795  GTGAGGAGTATTGG.....CCCTCCAAGCAGGCTCAGGACTATGGAGAC 3838
156  PheSerValSerCysGluAlaGluLysArgLysSerAspTyrIleIleAr 172
|||||:|||||:  |||:|||||:  |||:|||||:  |||:|||||:
3839  ATAAGTGGGCAATGACATCAGAAATTTCTTCGGAATGGACCATCAG 3888
172  gThrLeuLysValLys.....PheAsnSerGluThrArgThrIleTyrG 187
|  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
3889  AGATTTTCACAGTGAAAAATATCCACACAGTGAGAGTCAACCTCTGAGAC 3938
187  InPheHisTyrLysAsnTrpProAspHisAspValProSerSerIleAsp 203
|||||:|||||:  |||:|||||:  |||:|||||:  |||:|||||:
3939  AGTTCATTTTCACTCTCCGCGCAGACACGGTGTTCGCGACACCATGAC 3988
204  ProIleLeuGluLeuIleTrpAspValArgCysTyr.....GlnGluAs 218
|||||:|||||:  |||:|||||:  |||:|||||:  |||:|||||:
3989  CTGCTCATCAACTTCCGGTACTCTGTTGACTACATGACGACAGAGATCC 4038
218  pAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgThrG 235
|  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
4039  TCCCGAATCGCGATCTGGTGCATTCAGTGTGGGTGCGGAGGACGG 4088
235  lyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAspGly 251
|||||:|||||:  |||:|||||:  |||:|||||:  |||:|||||:
4089  GCATTTTCATTGCCATTGATCGTCTCATCTACCATAGAGAAATGAGAAC 4138
252  IleIleProGluAsnPheSerValPheSerLeuIleArgGluMetArgTh 268
|  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
4139  ACCGTG.....GATGTGTATGGGATGTGTATGACCTTCGAAT 4176
268  rGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValTyrA 285
|||||:|||||:  |||:|||||:  |||:|||||:  |||:|||||:
4177  GCATAGGCTTTAATGGTGCAGACAGAGACCATGATGTTTCTCCTCAATC 4226
285  snAlaValLeuGluLeuPheLysArgGlnMetAsp 296
|||||:|||||:  |||:|||||:  |||:|||||:
4227  AGTGTGTTTGGATATTGTGATGCCAGAAAGAC 4261

seq_name: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:PCT-US95-05512-1

seq_documentation_block:
; Sequence 1, Application PC/TUS9505512
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas K. and stman, Arne
; TITLE OF INVENTION: Density Enhanced Protein Tyrosine
; NUMBER OF INVENTION: Phosphatase
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; STREET: 233 South Wacker Drive, Suite 6300
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05512
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27866/31954
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5117 base pairs

```



```
US-09-600-358A-4 x US-08-018-129-4
Align seg 1/1 to: US-08-018-129-4 from: 1 to: 2790

14 AlaGlnSerLysLysIleThrLysGluGluPheAlaAsnGluPheLeu 30
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
836 GCTGAGACCAAGCAAGTAAAGTCAAAAGCGCTTTTGGGAAGAAATTTGAGAC 887
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
30 sLeuLysArgGlnSerThrLysLysLysAlaAspLysThrTyrProThr 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
888 ACTACAAACACAGAGTCAAACTT.....CTCTACAGCCGAA 925
47 hrValAlaGluAsnAlaLysAsnIleLysLysAsnArgTyrLysAsp 63
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
936 AAGAGGTCAAAGCGCAAGCAAAACAAAAATAGATATAAAACATC 975
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
64 LeuProTyrAspTyrSerArgValGluLeuSerLeuIleThrSerAsp 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
976 CTGCCCTTTGATCATACACAGGTTGTCTCACGATCTCTGATCCCAATGA 1025
80 uAsp...SerSerTyrIleAsnAlaAsnPheIleLysGlyValTyr.... 94
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1026 GCCTGTTTCAGATACATCAATGCAATATCATCATGCTCAATTTGAAA 1075
95 .....GlyProLysAlaTyrIleAlaThrGlnGly 104
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1076 CCAAGTGCACAAATTCAAAGCCCAAAAGAGTTACATTTGCCACAAAGGC 1125
105 ProLeuSerThrThrLeuLeuAspPheTrpArgMetIleTrpGluTyrSe 121
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1126 TGCCTGCAAAACACGCTGAATGACTTTTGGCGGATGTTCCAAAGAAA 1175
121 rValLeuIleIleValMetAlaCysMetGluTyrGluMetGlyLysLys 138
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1176 CTCGCGATGATGTTTCATGCAAGCAAAAGAGTGGAGAGAGAAAGATA 1225
138 yScyGluArgTyrTrpAlaGluProGlyGluMetGlnLeu...GluPhe 153
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1226 AATGTGCTCAATACTGG.....CTGTATGATGATGCTCTAAAGAATAT 1269
154 GlyProPheSerValSerCysGluAlaGluLysArgLysSerAspTyrI 170
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1270 GCGCTCATGCTGTAGGACGCTCAAGAAAGACGCGCTCATGACTATAC 1319
170 eIleArgThrLeuLysVal.....LysPheAsnSerGluThrA 183
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1320 CTAGAGAACTTAACCTTCAAGGTTGGACAGGAATACGGAG...A 1366
183 rgThrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValPro 199
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1367 GAACGGTCTGGCAATACCACTTTTCGGACCTGGCGGACCACGGCGTGCC 1416
200 SerSerIleAspProIleLeuGluLeuIleTrpValArgCysTyrGlu 216
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1417 AGCCACCTGGGGCGTCTGGACTTCCTGGAGGAGGTGCACCAATAGCA 1466
216 nGluAsp.....AspSerValProIleCysIleHisCysSerAlaGlyC 231
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1467 GGAGAGCATCATGATGACGGCGCGTGTGGTGCACGTGCAGTCTGGAA 1516
231 ysGlyArgThrGlyValIleCysAlaIleValAspTyrThrTrpMetLeu 247
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1517 TTGGCCGACAGGCGCTTC...ATTGATGATGATATTCATTATTCACATC 1563
248 LeuLysAspGlyIleIleProGluAsnPheSerValPheSerLeuIleAr 264
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1564 ATCAGAGAAAGGTGTGACTGCGATATTCACGTTCACCAAAACCATCCA 1613
264 gGluMetArgThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrG 281
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1614 GATGTGGGGTCTCAGAGTTCAGGGATGGTCCACAGAACACAGATACC 1663
281 luLeuValTyrAsnAlaVal.....LeuGluLeuPheLysArgGln 294
```

```
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1664 GATTTATCTATATGCGGTCAGCATTTATATTGAAACACTACAGCGCAGG 1713
295 MetAsp..... 296
1714 ATTGAAGAGAGACAGAAAAGCAAGAGGAAAGGCGACGAATATACAAATAT 1763
297 .....ValIleArgAspLysHisSerGlyThrGluSerGlnAlaLysH 311
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1764 TAAGTATTCTCTAGCGGACAGACGAGTGGAGATCAGAGCCCTCTCCCGC 1813
311 sCysIleProGlu..... 315
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1814 CTGTGACTCTTCGCGACCCCTGTGCAGAAATCAGAGAAAGACAGTCTAGA 1863
316 .....LysAsnHisThrLeuGlnAlaAspSerTyrSer.....Pr 327
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1864 GTCTATGAAAACGTGGCGCTGATGCAACAGCAGAAAAGTTTTCAGATGAGA 1913
327 oAsnLeuProLys.....SerThrThrLysAlaA 337
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1914 AAACCTGCCAAAACCTTCAGCAGCAAAATAGATGTGGACTTTCACCTCTCC 1963
337 laLysMetMetAsnGlnGlnArgThrLysMetGluIleLysGluSerSer 353
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1964 CTAAAAGATCAGGACGACGCAAGAAAGTTTATGTGAAGTCTGAATTT 2013
354 SerPheAsp.....PheArgThrSerGluIleSerAlaLysGlu 366
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2014 GGATTTGGAAGCGTTGCAATGTGGTTGACTACCTTTTGTATAAGCAAAAT 2063
366 uGluLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuG 383
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2064 TGAACCACTTAAAGACCACTGATTTTAACCTCAACAATACCTGCTCTCC 2113
383 luLeuAsnTyrSerPheAspLys 390
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2114 AATTACTCATTTCTCCAGATAAG 2136

seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-448-250-4

seq_documentation_block:
; Sequence 4, Application US/08448250
; Patent No. 5981251
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: Vogel, Wolfgang
; TITLE OF INVENTION: PTP ID: A NOVEL PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,250
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/018,129
; FILING DATE: 16-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-017
```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2790 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 130..1911
; US-08-448-250-4

alignment_scores:
  Quality: 496.00      Length: 441
  Ratio: 1.968         Gaps: 15
  Percent Similarity: 57.143      Percent Identity: 31.293

alignment_block:
  US-09-600-358a-4 x US-08-448-250-4 ..
  Align seq 1/1 to: US-08-448-250-4 from: 1 to: 2790

14  AlaglnSerLysLysIleThrLysGluGluPheAlaAsnGluPheLeuLy 30
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
838  GCTGAGACCCACAGATTAAGTCAACAAAGGCTTTGGGAAGAATTTGAGAC 887
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
30  sLeuLysArgGlnSerThrLysTyrLysAlaAspLysThrTyrProThr 47
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
888  ACTACAAACACAGAGTGCAACTT.....CTCTACAGCGCAA 925
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
47  hrValAlaGluAsnAlaLysAsnIleLysLysAsnArgTyrLysAspIle 63
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
926  AAGAGGGTCAAGGCAAGCAAAACAAACAAACAAATAGATATAAAACATC 975
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
64  LeuProTyrAspTyrSerArgValGluLeuSerLeuIleThrSerAspG1 80
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
976  CTGCCCTTTTCATCATCCAGGGTTGCTCCACGATCTGTGATCCCAATGA 1025
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
80  uAsp...SerSerTyrIleAsnAlaAsnPheIleLysGlyValTyr.... 94
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1026  GCCTGTTTCAGATTACATCAATGCAATATATCATCGCTGGAATTTGAAA 1075
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
95  .....GlyProLysAlaTyrIleAlaThrGlnGly 104
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1076  CCAAGTGCAACAATTCAAAGCCCAAAAGAGTTACATTGCCACACAAGGC 1125
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
105  ProLeuSerThrThrLeuLeuAspPheTrpArgMetIleThrGluTyrSe 121
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1126  TGCCTGCAAAACACGGTGAATGACTTTTGGCGGATGGTGTCCAAGAAA 1175
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
121  rValLeuIleIleValMetAlaCysMetGluTyrGluMetGlyLysLysL 138
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1176  CTCCCGAGTGATGTCATGACAGCAAGAAAGTGGAGAGGAAAGAGTA 1225
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
138  ysCysGluArgTyrTrpAlaGluProGlyGluMetGlnLeu...GluPhe 153
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1226  AATGTGCTCAATACTGG.....CCTGATGAGTATGCTCTAAAGAATAT 1269
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
154  GlyProPheSerValSerCysGluAlaGluLysArgLysSerAspTyrII 170
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1270  GCGGTCATGCGTGTAGGACGCTCAAGAAAGCGCCGCTCATGACTATAC 1319
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
170  eileArgThrLeuLysVal.....LysPheAsnSerGluThrA 183
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1320  GCTAAGAGAACTTAACCTTCAAGGTTGGACAAGGAATACGGAG...A 1366
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
183  rgThrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValPro 199
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

1367  GAACGGCTCTGGCAANTACCACCTTTCGGACCTGGCCGGACACCGCGGTGCC 1416
200  SerSerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrG1 216
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1417  AGCGACCTGGGGCGTGTGGACTTCCTGGAGGAGTGCACCATTAAGCA 1466
216  nGluAsp.....AspSerValProIleCysIleHisCysSerAlaGlyC 231
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1467  GGAGAGCATCATGATGCAGGGCGCGTGTGGTGCACACTGCAGTCTGGAA 1516
231  ysGlyArgThrGlyValIleCysAlaIleValAspTyrThrTrpMetLeu 247
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1517  TTGGCCGACAGGGACGCTTC...ATTGTGATGATATCTTATTATGACAT 1563
248  LeuLysAspGlyIleIleProGluAsnPheSerValPheSerLeuIleAr 264
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1564  ATCAGAGAGAAAGGTGTGACTCCGATATTGACGTTCCCAAAACCATCCA 1613
264  gGluMetArgThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrG 281
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1614  GATGGTGGGTCTCTCAGAGGTCAGGATGGTCCAGACAGAGAAGCACAGTAC 1663
281  lLeuValTyrAsnAlaVal.....LeuGluLeuPheLysArgGln 294
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1664  GATTTATCTATATGGCGTCCACGATTTATATTGAAACACTACAGCGCAGG 1713
295  MetAsp..... 296
1714  ATTGAAGAGAGCAGAAAGCAAGCAAGAAAGGCGCAGCAATATACAATAT 1763
297  .....ValIleArgAspLysHisSerGlyThrGluSerGlnAlaLysH 311
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1764  TAAGTATCTCTAGCGACACGAGTGGAGATCAGAGCCCTCTCCCGC 1813
311  iScysIleProGlu..... 315
1814  CTTGTACTCTCTCCGCCACCCCTGTGCAGAAATAGAGAGAAGACAGTGTAG 1863
316  .....LysAsnHisThrLeuGlnAlaAspSerTyrSer.....Pr 327
1864  GTCATGAAAGAGTGGCGCTGTATGCACACAGCAGCAAGAAAGTTTCAGAT 1913
327  oAsnLeuProLys.....SerThrThrLysAlaA 337
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1914  AAACCTGCCAAAACCTTCAGCACAGAAATAGATGTGGACTTTTCACCTCC 1963
337  lalysMetMetAsnGlnArgThrLysMetGluIleLysGluSerSer 353
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1964  CTAAGAGATCAGGACCAAGCAAGAAAGTTATGTGAAGTCTGAATTT 2013
354  SerPheAsp.....PheArgThrSerGluIleSerAlaLysG1 366
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2014  GGATTTGGAAGCGTTGCAATGTGTGACTACCTTTTGTGATAAGCAAAATT 2063
366  uGluLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuG 383
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2064  TGAACCATTTAAAGACCACTGTATTTTAACCAACAATACCTGCTTCCC 2113
383  lLeuAsnTyrSerPheAspLys 390
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2114  AATTACTCATTTCTCAGATAAG 2136
seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-449-644-4
seq_documentation_block:
; Sequence 4, Application US/08449644
; Patent No. 5856162
; GENERAL INFORMATION:
; APPLICANT: Schlessinger, Joseph
; APPLICANT: Sap, Jan M.
; APPLICANT: Ullrich, Axel
; APPLICANT: Vogel, Wolfgang
; APPLICANT: Fuchs, Miriam

```



```

3733 .....ATCGTCACACAATACCCCTCTGCCAAACACTGTAAAA 3768
      ::::: ::::: :::::
380 AspPheLeuGluLeuAsnTyrSerPhe..... 388
      ::::: ::::: :::::
3769 GACTCTCGAGATTAGTGTATGATGCTGTACCTCCATTGTGATGTT 3818
      ::::: ::::: :::::
389 .....AspLysAsnAlaAspThrThrMetLysTrpGlnThrLysA 402
      ::::: ::::: :::::
3819 AAACGAAGTCGACTGTCCACGGCTGCCCTCAGTACTGCCAGAGGAG 3868
      ::::: ::::: :::::
402 laPhe.....ProfileValGlyGluProLeuGlnLysHisGln 414
      ::::: ::::: :::::
3869 GGATGCTACGATAGGCCCCCAAGTGAATGATGCTCTGTTCAATG 3918
      ::::: ::::: :::::
415 SerLeuAspLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysPr 431
      ::::: ::::: :::::
3919 GACTGTGATGTGATCAACCGGATTTTAGGATATGCAATCTAACAGACC 3968
      ::::: ::::: :::::
431 O 431
3969 A 3969

seq_name: /cgn2_6/ptodata/1/1na/5B_COMB.seq:US-08-087-244A-4
seq_documentation_block:
; Sequence 4, Application US/08087244A
; Patent No. 5863755
; GENERAL INFORMATION:
; APPLICANT: Schlessinger, Joseph
; APPLICANT: Sap, Jan M.
; APPLICANT: Ullrich, Axel
; APPLICANT: Vogel, Wolfgang
; APPLICANT: Fuchs, Miriam
; APPLICANT: Fuchs, Miriam
; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASE-KAPPA
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,244A
; FILING DATE: 01-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-042
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4651 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4317
US-08-087-244A-4
```

```

alignment_scores:
  Quality: 492.50      Length: 467
  Ratio: 1.954        Gaps: 16
  Percent Similarity: 53.961  Percent Identity: 29.122

alignment_block:
US-09-600-358A-4 x US-08-087-244A-4 ..
Align seg 1/1 to: US-08-087-244A-4 from: 1 to: 4651

21  LysGluGluPheAlaAsnGluPheLeuLysLeuLysArgGlnSerThrLy 37
    |||||::: ::::: |||
2662 AAAGAGGAATATGAGAGCTTTT .....GAAGGACAGTCAGCATC 2702
    ::::: ::::: :::::

37  sTyrLysAlaAspLysThrTyrProThrThrValAlaGluAsnAlaLysA 54
    ::::: ::::: :::::
2703 TTGG .....GATGTAGCTAAAAAAGATCAAA 2728
    ::::: ::::: :::::

54  snlleLysLysAsnArgTyrLysAspIleLeuProTyrAspTyrSerArg 70
    |||||::: ::::: |||
2729 ATAGAGCAAAACCGATATGGAACATATATAGCATATGATCCTCCAGA 2778
    ::::: ::::: :::::

71  ValGluLeuSerLeuIleThrSerAspGluAspSerSerTyrIleAsnAl 87
    |||||::: ::::: |||
2779 GTGATTTTGCACACCGTAGAGGATGATCTCTCCTCAGATTATATTATGC 2828
    ::::: ::::: :::::

87  aasnPheIleLysGlyValTyrGlyProLysAlaTyrIleAlaThrGlnG 104
    |||||::: ::::: |||
2829 CAACTATATATGATGGCTACAGAGACCAAGTCATTACATTTGCCAACCAAG 2878
    ::::: ::::: :::::

104  LyProLeuSerThrThrLeuLeuAspPheTrpArgMetIleTrpGluTyr 120
    |||||::: ::::: |||
2879 GTCCCGTTCATGAACAGTGATGATTTCTGGAGGATGATTTGGCAAGAA 2928
    ::::: ::::: :::::

121  SerValLeuIleIleValMetAlaCysMetGluTyrGluMetGlyLysLy 137
    ::::: ::::: :::::
2929 CAATCTGCTTGCAATGTTGATGTTACAAATTTAGTTGAGTTGGCGGGT 2978
    ::::: ::::: :::::

137  sLysCysGluArgTyrTrpAlaGluProGlyGluMetGlnLeuGluPheG 154
    |||||::: ::::: |||
2979 TAAATGCTATAAATATTTGGCCTGATGATGACTGAAGTT .....TATG 3019
    ::::: ::::: :::::

154  LyProPheSerValSerCysGluAlaGluLysArgLysSerAspTyrIle 170
    |||||::: ::::: |||
3020 GTGACTTCAAGTAACTGTTAGAAATGGAACCACTTCTGCTGAATATGTA 3069
    ::::: ::::: :::::

171  IleArg.....ThrLeuLysValLysPheAsnSerGluThrArgThrI 185
    ::::: ::::: :::::
3070 GTTAGACATTCCACCTCGAAAGAGGGGTACAAATGAAATCCGTGAAGT 3119
    ::::: ::::: :::::

185  eTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSerSerI 202
    ::::: ::::: :::::
3120 TAAACAGTTCCATTTACCGGCTGGCTGACCATGAGTGAGTCCCTACCATG 3169
    ::::: ::::: :::::

202  leAspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnGluAsp 218
    ::::: ::::: :::::
3170 CTACAGGGCTGCTTCTTCTTATCCGGCGAGTCAAGTTATCAAAACCTCC 3219
    ::::: ::::: :::::

219  AspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgThrGl 235
    ::::: ::::: :::::
3220 AGTGTGTGCCCATCGTTGATATTGAGTGTGCTGGTGGAGCAACTGG 3269
    ::::: ::::: :::::

235  yValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAspGlyI 252
    |||||::: ::::: |||
3270 CTGCTACATTTGATTTGACATCATGCTAGCATGGCTGAAGAGAGGGGTG 3319
    ::::: ::::: :::::

252  leIleProGluAsnPheSerValPheSerLeuIleArgGluMetArgThr 268
    ::::: ::::: :::::
3320 TTGTT .....GATATTACAATTGTCAAGCCTTAAGATCT 3357
    ::::: ::::: :::::

269  GlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValTyrAs 285
    ::::: ::::: :::::

```



```
109 ThrLeuLeuAspPheTrpArgMetIleTrpGluTyrSerValLeuIle1 125
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1135 ACAGTCAATGACTTCTGGCAGATGGCTGGCAGGAGAACAGCGTGTCAT 1184
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
125 eValMetAlaCysMetGluTyrGluMetGlyLysCysGluArgT 142
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1185 CGTCATGACCCCGAGAGGTGGAGAAAGCGCGACAAATCGTCCCAT 1234
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
142 yTrpAlaGluProGlyGluMetGluLeuGluPheGlyProPheSerVal 158
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1235 ATGGCCCCAGGTGGC...ATGCAGCTGGCTTATGGGCCCTACTCTGTG 1281
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
159 SerCysGluAlaGluLysArgLysSerAspTyrIleIleArgThrLeuLy 175
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1282 ACCAAGTCGGGGAGCATCACACACCGAATACAACTCCGTACCTTACA 1331
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
175 sVallys.....PheAsnSerGlu...ThrArgThrIleTyrGlnPheH 189
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1332 GGTCTCCCGCTGGACAATGGAGACCTGATTCGGGAGATCTGGCATTACC 1381
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
189 isTyrLysAsnTrpProAspHisAspValProSerSerIleAspProIle 205
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1382 AGTACCTGAGCTGGCCCGACCAATGGGTCCTCCAGTAGGCTGGGGGTGC 1431
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
206 LeuGluLeuIleTrpAspValArgCysTyrGlnGluAsp.....AspSe 220
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1432 CTCAGCTTCTGGACCATCAACACAGCGGAGGAAAGTGCCTCACGC 1481
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
220 rValProIleCysIleHisCysSerAlaGlyCysGlyArgThrGlyValI 237
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1482 AGGGCCCATCATCTGTCGACATGCGCGCGGCATCGCGCCGACAGGCACA 1531
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
237 leCysAlaIleValAspTyrThrTrpMetLeuLeuLysAspGlyIleIle 253
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1532 TCATTGTCTATC.....GACATGCTC 1551
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
254 ProGluAsnPheSerValPheSerLeu.....GATCTGCTC 262
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1552 ATGGAGAATCATCTCCACCAAGGCGCTGGACTGTGACATTGACATCCAGAA 1601
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
263 ...IleArgGluMetArgThrGlnArgProSerLeuValGlnThrGlnG 278
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1602 GACCATCCAGATGGTGGCGGCGCGCGCTCGGGCATGGTGGACAGCGGAG 1651
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
278 luGlnTyrGluLeuValTyrAsnAlaVal.....LeuGluLeuPhe 291
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1652 CGCAGTACAGTTCATCTACGTGGCCATCGCCCGCTGTTCAATGAACCACT 1701
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
292 LysArgGlnMetAspValIleArgAspLysHisSerGlyThrGluSerGl 308
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1702 AAGAGAAGCTGGAGTCTCTGCAGTCCGAG...AAGGGCCAGAGTCGGA 1748
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
308 nAlaLysHisCysIleProGluLysAsnHisThrLeuGlnAlaAspSerT 325
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1749 GTACGGGAATC.....ACCT 1765
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
325 yTrSerProAsnLeuProLysSerThrThrLysAlaAlaLys..... 338
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1766 ATCCCCACCCATGAAGATGCGCCATGCCAAGCCCTCCCGCCTCGTCC 1815
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
339 .....MetMetAsnGlnGlnArgThrLy 346
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1816 AAACACAAGGAGGATGTGTATGAGAACCTGCACACTAAGAACAAGAGGA 1865
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
346 smetGluIleLysGluSerSerPheAspPheArgThrSerGluIleS 363
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1866 GGAGAAGTGAAGAACGCGTCTACAGACACAAAGGAGAGCAAGGGTT 1915
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
363 erAlaLysGluGlu 367
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1916 CCTCAAGAGGAAG 1929
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

seq_name: /cgn2_6/ptodata/1/1na/6A_COMB.seq:US-09-358-685-1

```
seq_documentation_block:
; Sequence 1, Application US/09358685
; Patent No. 6121047
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF SHP-1 EXPRESSION
; FILE REFERENCE: RTS-0081
; CURRENT APPLICATION NUMBER: US/09/358,685
; CURRENT FILING DATE: 1999-07-21
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 2277
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (256)..(2049)
US-09-358-685-1
```

alignment_scores:

Quality: 485.00 Length: 388
Ratio: 2.012 Gaps: 14
Percent Similarity: 62.113 Percent Identity: 32.990

alignment_block:

US-09-600-358A-4 x US-09-358-685-1 ..

Align seg 1/1 to: US-09-358-685-1 from: 1 to: 2277

```
15 GlnSerLysLysIleThrLysGluGluPheAlaAsnGluPheLeuLysLe 31
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
964 GAGTCCGAGGATACAGCAAGCGTGGCTTCGGAGGAGTTTGAGAGTTT 1013
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
31 uLysArgGlnSerThrLysTyrLysAlaAspLysThrTyrProThrThrv 48
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1014 GCAGAAGCAGGAGGTCAAG.....AACTTGACACGAGCGTCTGG 1051
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
48 alAlaGluAsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeu 64
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1052 AAGGCGACGCGCCAGAGAAACAGGGCAAGACCGCTACAAGAACATCTC 1101
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
65 ProTyrAspTyrSerArgValGluLeuSerLeuIleThrSerAsp...Gl 80
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1102 CCCTTTGACCACGCGAGTGTCTTCGAGGACGGGACAGTAACATCCC 1151
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
80 uAspSerSerTyrIleAsnAlaAsnPheIleLysGly...ValTyrGlyP 96
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1152 CGGGTCCGACTACATCAATGCCAACTACATCAAGAACACGCTGCTAGGCC 1201
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
96 ro.....LysAlaTyrIleAlaThrGlnGlyProLeuSerThr 108
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1202 CTGATGAGAACGCTAAGACCTACATCGCAGGACGGGCTGTCTGGAGGCC 1251
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
109 ThrLeuLeuAspPheTrpArgMetIleTrpGluTyrSerValLeuIle1 125
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1252 ACGGTCAATGACTTCTGGCAGATGGCTGGCAGGAGAACAGCGGTGTCAT 1301
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
125 eValMetAlaCysMetGluTyrGluMetGlyLysLysLysCysGluArgT 142
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1302 CGTCATGACCCCGAGAGGTGGAGAAAGCGCGAACAATCGTCCCAT 1351
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
142 yTrpAlaGluProGlyGluMetGluLeuGluPheGlyProPheSerVal 158
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1352 ACTGGCCCCAGGTGGGC...ATGCAGCGTGTATGGGCCCTACTCTGTG 1398
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
159 SerCysGluAlaGluLysArgLysSerAspTyrIleIleArgThrLeuLy 175
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1399 ACCAACTGCGGGAGCATGACACAAACCGGAATACAAACTCCGTACCTTACA 1448
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
175 sVallys.....PheAsnSerGlu...ThrArgThrIleTyrGlnPheH 189
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentLin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/202,389
FILING DATE: 28-FEB-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/983,926
FILING DATE: 01-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/829,141
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/721,112
FILING DATE: 26-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: BIH92-05MA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2143 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 145..2037
145-08-202-389-7

```
alignment_scores:
  Quality: 484.00      Length: 443
             Ratio: 1.847      Gaps: 18
  Percent Similarity: 59.142  Percent Identity: 31.603
```

```
alignment_block:
US-09-600-358A-4 x US-08-202-389-7 ..
```

Align seq 1/1 to: US-08-202-389-7 from: 1 to: 2143

15 GlnSerLysLysIleThrLysGluGlupheAlaAsnGlnupheLeuLysLe 31
:::||||: ::||| ||| ::||| ||| ::||
247 GAGTCCGAGGATACGCCAAGGCTGGCTTCCTGGGAGGAGCTTTGAGAGTTT 896

seq_name: /cqn2_6/ptodata/1/ina/5A_COMB.seq;US-08-202-389-7

seq documentation block:

; Sequence 7, Application US/08202389

; Patent No. 5536636

GENERAL INFORMATION:

APPLICANT: Freeman Jr., Robert M.

; APPLICANT: Plutzky, Jorge

; APPLICANT: Neel, Benjamin G.

; APPLICANT: Rosenberg, Robert D.

; TITLE OF INVENTION: IDENTIFICATION OF NOVEL TYROSINE

; TITLE OF INVENTION: PHOSPHATASES HAVING SH2 DOMAINS

; NUMBER OF SEQUENCES: 54

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Militia Drive

CITY: Lexington

STATE: MA

109 ThrLeuLeuAspPheTrpArgMetIleTrpGluTyrSerValLeuIle1 125
1135 ACGGTCAATGACTTCGGCAGATGGCGTGGCAGGAGAACAGCGGTGCAT 1184
125 eValMetAlaCysMetGluTyrGluMetGlyLysLysCysGluArgT 142
1185 CGTCATGACCCAGCGAGGTGGAGAAAGCGGGAACAAATGCGTCCCAT 1234
142 yTrpAlaGluProGlyGluMetGluLeuGluPheGlyProPheSerVal 158
1235 ACTGGCCCGAGGTGGC...ATGCACGTGCTTATGGCCCTACTCTGTG 1281
159 SerCysGluAlaGluLysArgLysSerAspTyrIleIleArgThrLeuLy 175
1282 ACCAACTGGGGAGGATGACACCAACCGAATACAACTCCGTACCTTACA 1331
175 sValLys.....PheAsnSerGlu...ThrArgThrIleTyrGlnPheH 189
1332 GGTCTCCCGCTGGACAATGGAGACCTGATCGGGAGATCTGGCAATTACC 1381
189 iStryLysAsnTrpProAspHisAspValProSerSerIleAspProIle 205
1382 AGTACCTGAGTGGCCCGACCATGGGGTCCCGAGTGGCGGTGGGTGTC 1431
206 LeuGluLeuIleTrpAspValArgCysTyrGlnGluAsp.....AspSe 220
1432 CTCAGCTTCCTGGACCATCAACACCGCGCAGGAAGTCTGCCTCACGC 1481
220 rValProIleCysIleHisCysSerAlaGlyCysGlyArgThrGlyValI 237
1482 AGGGCCCATCATCTGCACTGCAGCGCGCGCATCGCGCACAGGCACCA 1531
237 leCysAlaIleValAspTyrThrTrpMetLeuLeuLysAspGlyIleIle 253
1532 TCATTGTCATC.....GACATGCTC 1551
254 ProGluAsnPheSerValPheSerLeu..... 262
1552 ATGGAGAACATCTCCACCAAGGCGCTGGACTGTGACATTGACATCCAGAA 1601
263IleArgGluMetArgThrGlnArgProSerLeuValGlnThrGlnG 278
1602 GACCATCCAGATGGTGGCGCGCAGCGCTGGCATGGTGCAGACGGAGG 1651
278 luGlnTyrGluLeuValTyrAsnAlaVal.....LeuGluLeuPhe 291
1652 CGCAGTACAAAGTTCATCTACGTGGCCATCGCCCGCATGTTCAATGAACCACT 1701
292 LysArgGlnMetAspValIleArgAspLysHisSerGlyThrGluSerG1 308
1702 AAGAAGAAGCTGGAGTCTCTGCTGCTGCGAG...AAGGCCAGAGTCCGA 1748
308 nAlaLysHisCysIleProGluLysAsnHisThrLeuGlnAlaAspSerT 325
1749 GTACGGGAACATC.....ACCT 1765
325 yrSerProAsnLeuProLysSerThrThrLysAlaAlaLysMetMetAsn 341
1766 ATCCCCGAGCATGAGATGCGCCATGCCAAGGCCTCC..... 1803
342 GlnGlnArgThrLysMetGluIleLysGluSerSerSerPheAspPheAr 358
1804CGCACCCTCGTCCAAACACACAGGAGGATGTGTATCAGAACCTGCA 1847
358 gThrSerGluIleSerAlaLysGluGluLeuValLeuHisProAlaLys 375
1848 CACTAAG.....AACAGAGGAGGAA..... 1869
375 erSerThrSerPheAspPheLeuGluLeuAsnTyrSerPheAspLysAsn 391
1870AGTGAAGAA 1878

392 AlaAspThrThrMetLysTrpGlnThrLysAlaPheProIleValGlyG1 408
1879 GCAGCGGTGACAGACAGGAGAGCAAGAGGTTCCCT..... 1917
408 uProLeuGlnLysHisGlnSerLeuAspLeuGlySerLeuLeuPheGluG 425
1918CAAGAGGAAGTAGC.....GGTGTGTCTTCAGGTGGC 1951
425 lYcysSerAsnSerLysProValAsnAla 434
1952 CATGCTCAGCCCTGACCTGTGGAGCA 1980
seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-202-389-11
seq_documentation_block:
; Sequence 11, Application US/08202389
; Patent No. 5536636
; GENERAL INFORMATION:
; APPLICANT: Freeman Jr., Robert M.
; APPLICANT: Plutzky, Jorge
; APPLICANT: Neel, Benjamin G.
; APPLICANT: Rosenberg, Robert D.
; TITLE OF INVENTION: IDENTIFICATION OF NOVEL TYROSINE
; TITLE OF INVENTION: PHOSPHATASES HAVING SH2 DOMAINS
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/202,389
; FILING DATE: 28-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/983,926
; FILING DATE: 01-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/829,141
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/721,112
; FILING DATE: 26-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: B1H92-05MA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2276 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 114..1893
US-08-202-389-11

alignment_scores:
Quality: 483.00 Length: 488
Ratio: 1.851 Gaps: 20

```

1586 CAAAACATCCAGATGGTGGGTCTCAGAGGTCAGGATGGTCCAGACAG 1633
277  InGluGlnTyrGluLeuValTyrAsnAlaVal.....LeuGluLeu 290
:: |||||:::||||| |||||:::|||||
1635 AAGCAGATCCGATTTATCTATATCGCGTCCAGCATTTATATTGAACA 1685
291 PheLysArgGlnMetAsp..... 296
:::||||:::|:::|
1686 CTACAGCGCAGGATTCAAGAAAGACGAGAAAAGCAAGAGGAAGGCGCACGA 1735
297 .....ValIleArgAspLysHisSerGlyThrGluS 307
::: |||||:::|||||
1736 ATATACAAATATTAACTATCTCTACGCGACGACGAGTGGAGATCAGA 1785
307 eGlnAlaLysHisCysIleProGluLysAsnHisThrLeuGlnAlaAsp 323
|| || ||| |||
1786 GCCCTCTCCGCTTGACTCCA..... 1808
324 SerTyrSerProAsnLeuProLysSerThrThrLysAlaAlaLysMetMe 340
||| |||:::|||||
1809 .....ACGCCACCTGTGCAAAATGAG 1831
340 tAsnGlnGlnArgThrLysMetGlu.....IleLysGluSerS 353
:::||||| |||
1832 AGAAGACAGATGCTAGAGTCTATGAAAACCTGGCCCTGATGCAACACGAGA 1881
353 eSerPhe AspPheArgThrSerGlu...IleSerAlaLysGluGluLe 368
:::||||| ||| |||||:::|||||
1882 AAGTTTCAGATGAGAAAACCTGCCAAAACCTTCAGCACAGAAATAGATGT 1931
368 uValLeuHisPro.....AlaLysSerSerThrSerPheAspp 381
:::||||| ||| |||||:::
1932 GGACTTTCACCTCTCCCTAAAAAGATCAAGAACAGACGCAAGAAAGTTT 1981
381 heLeuGluLeuAsnTyrSerPheAspLysAsnAlaAspThrThrMetLys 397
:::|||||:::|||||::: |||
1982 ATGTGAAGACAGAAATTTGGATTTGGAAGCGCTTGCA.....ATG 2019
398 TrpGlnThr.....LysAlaPheProIleValGlyGluProLe 410
||| ||| |||:::|||||
2020 TGGTTGACTACCTTTTGATTAAGCAAATTTGAACCATTTAAAGACCACCT 2069
410 uGlnLysHisGlnSerLeuAspLeuGlySerLeuLeuPheGluGlyCys 427
| .....
2070 G..... 2070
444 ProIleThrArg 447
|||:::|
2089 CTGCTTCCCAA 2100

seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-449-644-3
seq_documentation_block:
; Sequence 3, Application US/08449644
; Patent No. 5856162
; GENERAL INFORMATION:
; APPLICANT: Schlusser, Joseph
; APPLICANT: Sap, Jan M.
; APPLICANT: Ullrich, Axel
; APPLICANT: Vogel, Wolfgang
; APPLICANT: Fuchs, Miriam
; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASE-KAPPA
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York

```

```

? COUNTRY: U.S.A.
? ZIP: 10036
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/449,644
? FILING DATE: 24-MAY-1995
? CLASSIFICATION: 514
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/087,244
? FILING DATE: 01-JUL-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Coruzzi, Laura A.
? REGISTRATION NUMBER: 30,742
? REFERENCE/DOCKET NUMBER: 7683-042
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 212-790-9090
? TELEFAX: 212-869-8864/9741
? TELEX: 66141 PENNIE
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 4374 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: unknown
? MOLECULE TYPE: DNA (genomic)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1..4371
? US-08-449-644-3

alignment_scores
  Quality: 478.50      Length: 473
  Ratio: 1.899        Gaps: 17
  Percent Similarity: 53.277      Percent Identity: 28.541

alignment_block
  us-09-600-358a-4 x us-08-449-644-3
  Align seg 1/1 to: us-08-449-644-3 from: 1 to: 4374

21  LysGluGluPheAlaAsnGluPheLeuLysLeuLysArgGlnSerThrLy 37
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2698  AAGAGGAATACGAGAGCTCTTT.....GAAGGCCAGTCAGCCTC 2738

37  sTyrLysAlaAspLysThrTyrProThrValAlaGluAsnAlaLysA 54
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2739  TTGG.....GATGTGGCTAAAGAGATCAAA 2764

54  snlleLysLysAsnArgTyrLysAspIleLeuProTyrAspTyrSerArg 70
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2765  ACAGAGCAAGAACCGATACGGAACATTTATCGCATATGATCACTCCAGA 2814

71  ValGluLeuSerLeuIleThrSerAspGluAspSerTyrIleAsnAl 87
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2815  GTCATCTGCAACCTGTGGAAGATGACCTCTTCAGATTACATTATATGC 2864

87  aAsnPhelle.....LysGlyValTyrGlyProLysA 98
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2865  CAACATACATGACATTTGGTGTACAGGATGGCTACCGAGACCAAGCC 2914

98  latyrIleAlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrp 114
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2915  ACTACATGCAACTCAAGCCAGCTTCATGAACCGTATATGATTTTGG 2964

115  ArgMetIleTrpGluTyrSerValLeuIleValMetAlaCysMetGl 131
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2965  AGGATGTGTGGCAAGACGAGCTGCTGTATTGTGTGTCACATAATTT 3014

```

```

131  uTyrGluMetGlyLysLysCysGluArgTyrTrpAlaGluProGlyG 148
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3015  AGTGGGAAGTGGCGGCTAAATGCTATAAATATTGGCCTGATGATAC 3064

148  luMetGlnLeuGluPheGlyProPheSerValSerCysGluAlaGluLys 164
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3065  AGGTT.....TATGGTGACTTCAAAGTCACTCGCTAGAAATGGAG 3105

165  ArgLysSerAspTyrIleIleArg.....ThrLeuLysValLysPheAs 179
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3106  CCACCTTGCTGAGTATGTCGTAGGACATTCACCTTGGAAAGGAGGGCTA 3155

179  nSerGluThrArgThrIleTyrGlnPheHisTyrLysAsnTrpProAspH 196
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3156  TAATGAATCCGTTGAAGTCAACAGTTCACCTTCACTGGCTGGCTGACC 3205

196  isAspValProSerSerIleAspProIleLeuLeuIleTrpAspVal 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3206  ATGGTGTTCATACCAACGCAACAGGCTCTCTGTCATTTATCCGGAGAGTC 3255

213  ArgCysTyrGlnGluAspAspSerValProIleCysIleHisCysSerAl 229
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3256  AAGCTATCTAACCCCTCCAGTGTGGGCCCATTTGCTACACATGCAGTGC 3305

229  aGlyCysGlyArgThrGlyValIleCysAlaIleValAspTyrThrTrpM 246
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3306  TGGTGTGGCGCACAGGCTGTACATTGTTATTGACATAATGCTGGACA 3355

246  etLeuLeuLysAspGlyIleIleProGluAsnPheSerValPheSerLeu 262
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3356  TGGCTGAAAGAGAGGGTGTGGTT.....GACATCTACAACCTGT 3393

263  IleArgGluMetArgThrGlnArgProSerLeuValGlnThrGlnGluGl 279
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3394  GTGAAAGCCCTTACGATCTCGGCCATTAATATGGTACAGACAGAGGAACA 3443

279  nTyrGluLeuValTyrAsnAlaValLeuGluLeu..... 290
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3444  GTACATTTTATTCATGATGCCCATTTTAGAAGCCTGCTTATGTGGAGAAA 3493

291  .....PheLys....ArgGlnMetAspValIleArg 299
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3494  CTGCCATCCCTGTGTGTGAATTTAAAGCTGCATATTTTGATATGATCGA 3543

300  .....AspLysHisSerGlyThrGluSerGlnAlaLys 310
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3544  ATGACTCTCAGACTAACTCTCTCATCTCAAAGATGAATTCAGACTCT 3593

310  sHis.....CysI 313
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3594  GAATTCGGTCAACCCCTCGACTCAAGCTGAAGACTGCAGCATAGCCTGCC 3643

313  leProGluLysAsnHisThrLeuGln.....AlaAspSerTyrSerPro 327
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3644  TGCCA...AGGAACCATGACAAGAACCGTTTCATGGATATGCTCCACCT 3690

328  Asn.....LeuProLysSerThrThrLysAlaAlaLysMetMetAsnGl 342
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3691  GACAGATGCTGCTGCTTTTAAATTAATCAATTTGATGGGGAGAGCAGTAACTA 3740

342  nGlnArgThrLysMet.....GluIleLysGluSerSerPheAspP 357
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3741  CATCAATGCTGCTTTATGATAGCTATAGCAGCCAGCAGCTTC.... 3786

357  heArgThrSerGluIleSerAlaLysGluLeuValLeuHisProAla 373
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3787  .....ATCGTCAACAATACCCA 3804

374  LysSerSerThrSerPheAspPheLeuGluLeuAsnTyrSerPhe..... 388
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3805  CTGCCAAACACTGTGAAAGACTTCTGGAGATTAGTATATGATTCAGGATG 3854

389  .....AspLysAsnAlaAspThrThrM 396

```

```

3855 TACCTCCATCGTGATGCTAAATGAAGTGGACCTGTCTCAGGCTGCCAC 3904
|||
396 etLysTrpGlnThrLysAlaPhe.....ProIleValGlyGlu 408
|||||
3905 AGTACTGCCGACAGAAAGAAATGCTGCAATGCTTATCCAAAGTGGA 3954
|||||
409 ProLeuGlnLysHisGlnSerLeuAspLeuGlySerLeuLeuPheGluG1 425
|||||
3955 TGTATGCTTCTTCAATGGACGTGTGATGTGATCAATCCAAATTTTAGAAT 4004
|||||
425 yCysSerAsnSerLysPro 431
|||||
4005 ATGCAACCTAACGAGACCA 4023
seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-087-244A-3
seq_documentation_block:
; Sequence 3, Application US/08087244A
; Patent No. 5863755
; GENERAL INFORMATION:
; APPLICANT: Schlusser, Joseph
; APPLICANT: Sap, Jan M.
; APPLICANT: Ullrich, Axel
; APPLICANT: Vogel, Wolfgang
; APPLICANT: Fuchs, Miriam
; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASE-KAPPA
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,244A
; FILING DATE: 01-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-042
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4374 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4371
; US-08-087-244A-3

```

```

alignment_scores:
Quality: 478.50 Length: 473
Ratio: 1.899 Gaps: 17
Percent Similarity: 53.277 Percent Identity: 28.541
alignment_block:
US-09-600-358A-4 x US-08-087-244A-3

```

```

Align seg 1/1 to: US-08-087-244A-3 from: 1 to: 4374
21 LysGluGluPheAlaAsnGluPheLeuLysLeuLysArgGlnSerThrLy 37
|||||
2698 AAGAGGAATACGAGAGCTCTTT.....GAAGCCAGTCAGGCTC 2738
|||||
37 sTyrLysAlaAspLysThrTyrProThrThrValAlaGluAsnAlaLysA 54
|||||
2739 TTGG.....GATGTGGCTAAAAAGGATCAAA 2764
|||||
54 snlleLysLysAsnArgTyrLysAspIleLeuProTyrAspTyrSerArg 70
|||||
2765 ACAGAGCAAGAACGATACGGAACATATTCGCATATGATCACTCCAGA 2814
|||||
71 ValGluLeuSerLeuIleThrSerAspGluAspSerSerTyrIleAsnAl 87
|||||
2815 GTCATCTGCAACCTGTGGAAGATGACCTTCTTCAGATATACATTAATGC 2864
|||||
87 aasnPheIle.....LysGlyValTyrGlyProLysA 98
|||||
2865 CACTACATCGACATTTGGCTGTACAGGGATGGCTACCAGACACCAAGCC 2914
|||||
98 latyrIleAlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrp 114
|||||
2915 ACTACATTCGAACCAAGGCCAGTTCATGAACCGTATATGATTTTGG 2964
|||||
115 ArgMetIleTrpGluTyrSerValLeuIleValMetAlaCysMetG1 131
|||||
2965 AGGATGTGTGGCAAGACGACGCTGCTGATTTGATGGTCACTAATTT 3014
|||||
131 utyrGluMetGlyLysLysCysGluArgTyrTrpAlaGluProGlyG 148
|||||
3015 ACTGGAAGTTGGCGGTGAATGCTATAATATTTGGCTGATGATACTG 3064
|||||
148 luMetGlnLeuGluPheGlyProPheSerValSerCysGluAlaGluLys 164
|||||
3065 AGGTT.....TATGGTGACTTCAAGTCACCTGCTAGAAATGGAG 3105
|||||
165 ArgLysSerAspTyrIleIleArg.....ThrLeuLysValLysPheAs 179
|||||
3106 CCACCTGCTGATGATGCTTGAAGACATTCACCTTGGAAAGAGGGGCTA 3155
|||||
179 nserGluThrArgThrIleTyrGlnPheHisTyrLysAsnTrpProAspH 196
|||||
3156 TAATGAATCCGTGAAGTCAACAGTCCACTTCCTGCTGCTATTCGGAGAGTC 3205
|||||
196 isAspValProSerSerIleAspProIleLeuGluLeuIleTrpAspVal 212
|||||
3206 ATGGTGTTCATACCAACAGGCTCTGCTGCTATTCGGAGAGTC 3255
|||||
213 ArgCysTyrGlnGluAspAspSerValProIleCysIleHisCysSerAl 229
|||||
3256 AAGCTATCTAACCTCCAGTGGGCGCCATTCGTACACACTCAGTGC 3305
|||||
229 aGlyCysGlyArgThrGlyValIleCysAlaIleValAspTyrTrpM 246
|||||
3306 TGGTGTGGGCGCACAGGCTGTACATTTGATGACATATGCTGGACA 3355
|||||
246 etLeuLeuLysAspGlyIleIleProGluAsnPheSerValPheSerLeu 262
|||||
3356 TGGCTGAAAGAGAGGGTGTGGTT.....GACATCTACACTGT 3393
|||||
263 IleArgGluMetArgThrGlnArgProSerLeuValGlnThrGlnGluG1 279
|||||
3394 GTGAAGCCCTTACGATCTCGCGCATTAATATGTTACAGACAGAGGAACA 3443
|||||
279 nTyrGluLeuValTyrAsnAlaValLeuGluLeu..... 290
|||||
3444 GTACATTTTATTCATGATGCCATTTTAGAAGCCTGCTTATGTGGAGAAA 3493
|||||
291 .....Phelys...ArgGlnMetAspValIleArg 299
|||||

```

```
3494 CTGCCATCCCTGTGTGTGAATTTAAAGCTGCATATTTTGTATATGATTCGA 3543
300 .....AspLysHisSerGlyThrGluSerGlnAlaLy 310
      :::::||||| ||| |||:::
3544 ATAGACTCTCAGACTAACTCCTCTCATCTCAAGATGAATTCAGACTCT 3593
310 sHis.....CysI 313
      ::|
3594 GAATTCGGTCACCCCTCGACTACAAGCTGAAGACTGCAGCATAGCCTGCC 3643
313 leProGluLysAsnHisThrLeuGln.....AlaAspSerTyrSerPro 327
      ::||| ::|||::: ||| |||
3644 TGCCA...AGGAACCATGACAGAACCGTTTCATGGATATGCTCCACCT 3690
328 Asn.....LeuProLysSerThrThrLysAlaAlaLysMetMetAsnG1 342
      ::| ||||| ||| ::::: |||
3691 GACAGATGCTGCCTTTTAAATTAATGATGGGGAGAGCAGTAACCTA 3740
342 nGlnArgThrLysMet.....GluIleLysGluSerSerSerPheAspP 357
      ::::: ::| ::| ::| ::| ::| ::|
3741 CATCAATGCTGCTCTTATGGATAGCTATAGGCAGCCAGCAGCTTC.... 3786
357 heArgThrSerGluIleSerAlaLysGluGluLeuValLeuHisProAla 373
      ::||| ::|
3787 .....ATCGTCACACAATACCCA 3804
374 LysSerSerThrSerPheAspPheLeuGluLeuAsnTyrSerPhe..... 388
      ::||| ||||| ::||| |||:::
3805 CTGCCAAACACTGTGAAGACTTCTGGAGATTAGTATATGATTACGGATG 3854
389 .....AspLysAsnAlaAspThrThrm 396
      ||| ::|
3855 TACCTCCATCGTGTGCTAAATGAAGTGGACCTGTCTCAGGGCTGCCAC 3904
396 etLysTrpGlnThrLysAlaPhe.....ProIleValGlyGlu 408
      :: ||| ::::: |||||
3905 AGTACTGGCCAGAGAAGGAATGCTGCGATATGGTCTATCCCAAGTGGAA 3954
409 ProLeuGlnLysHisGlnSerLeuAspLeuGlySerLeuLeuPheGluG1 425
      ::::: ::::: |||::: ::| ::| ::| ::|
3955 TGTATGCTTGTTCATGGACTGTGATGTGATCAATCGAATTTTAGAAT 4004
425 yCysSerAsnSerLysPro 431
      |||::: ::::: |||
4005 ATGCAACCTAACGAGACCA 4023
```

[illegible]

```
alignment_scores:      Quality: 3615.00      Length: 692
                       Ratio: 5.224          Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-600-358A-4 x AAX90696 ..

Align seg 1/1 to: AAX90696 from: 1 to: 2356

1  MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17
|||||
42  ATGGACCAAGAGAAATCTCGAGAAGTTCTGGATGAGGCCCAAGCAA 91

17  sLysIleThrLysGluPheAlaAsnGluPheLeuLysLeuLysArgG 34
|||||
92  GAAATTTACTAAAGAGAGGTTTCCCAATGAATTTCTGAAGCTGAAAGGC 141

34  InSerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50
|||||
142  AATCTACCAAGTAGCAGGCGACAAACCTTATCTCAACTGTGGCTGAG 191

51  AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
|||||
192  AATGCCAAGAATATCAAGAAAAACAGATATAAGGATATTTGGCCCTATGA 241

67  pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSert 84
|||||
242  TTATAGCCGGGTAGAATATCCCTGATAACCTCTGATGAGGATTCACGCT 291

84  YrIleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
|||||
292  ACATCAATGCCAACTTCATTAAAGGAGTTTATGACCCCAAGGCTTATATT 341

101  AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetIl 117
|||||
342  GCCACCACCGGCTCTTATCTACAACCTCCTCGACTTCTGGAGGATGAT 391

117  eTrpGluTyr-SerValLeuIleIleValMetAlaCysMetGluTyrGluM 134
|||||
392  TTGGGAATATAGTGCTCTTATCATTTGTTATGGCATGCATGGATATGAAA 441

134  etGlyLysLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
|||||
442  TGGGAAGAAAAGTGTGAGGCTACTGGCTGAGCCAGAGAGATGCAG 491

151  LeuGluPheGlyPropheSerValSerCysGluAlaGluLysArgLysSe 167
|||||
492  CTGGAATTTGGCCCTTCTCTGTATCTCTGTAAGCTGAAAAAGGAAATC 541

167  rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
|||||
542  TGATATATATATAGGACTCTAAAAGTTAAGTTCAATAGTAGAACTCGAA 591

184  hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSer 200
|||||
592  CTATCTACCAAGTTTCATTACAAGAAATGGCCAGACCATGATGACCTTCA 641

201  SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnGl 217
|||||
642  TCTATAGACCTTATCTTGGCTCATCTGGGATCTAGCTTTGTACCAAGA 691

217  uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
|||||
692  GGATGACAGGTGTTCCCATATGCAATTCATCTACTGCAGTGGCTGGGAAGA 741

234  hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
|||||
742  CTGGTGTTATTTGGCTATTGTTGATTATACATGGATGTTGCTAAAAGAT 791

251  GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267
|||||
```

```
792  GGGATAAATTCCTGAGAACTTCAGTGTGTTTTTCAGTTTGCAGGAAATGCG 841

267  gThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValT 284
|||||
842  GACACAGAGGCCCTTCATTAGTTCAAACGCCAGGAACAATATGAACCTGTCT 891

284  yrAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
|||||
892  ACAATGCTGTATTAGAACTATTAAAGACAGATGGATGTTATCATCAGAT 941

301  LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
|||||
942  AAACATCTCTGGAACAGAGAGTCAAGCAAAAGCATTTGTTATCTCAGAAAAA 991

317  nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
|||||
992  TCACACTCTCCAAGCAGACTCTTATTCTCTTAATTTACCAAAAAGTACCA 1041

334  hrLysAlaAlaLysMetMetAsnGlnGlnArgThrLysMetGluIleLys 350
|||||
1042  CAAAACGACGCAAAATGATGAACCAACAAAGGACAAAAATGGAATCAAA 1091

351  GluSerSerSerPheAspPheArgThrSerGluIleSerAlaLysGluGl 367
|||||
1092  GAATCTCTCTCTTCTGACTTTAGGACTTCTGAAATAAGTGCAAAAGAAGA 1141

367  uLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGluL 384
|||||
1142  GCTAGTTTTTGCACCCCTGCTAAATCAAGCACTTCTTTGACTTTCTGGAGC 1191

384  euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnThr 400
|||||
1192  TAAATACAGTTTTGACAAAAATGCTGACACACCATGAATGCGACACA 1241

401  LysAlaPheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAs 417
|||||
1242  AAGGCATTTCCAATAGTTGGGAGCCCTCTTCAGAAGCATCAAAAGTTTGA 1291

417  pLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsnA 434
|||||
1292  TTTGGGCTCTCTTTGTTTGGGGATGTTCTTAATCTTAAACCTGTAATG 1341

434  laLaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSer 450
|||||
1342  CACGAGGAAGATATTTAATTCAAAGGTGCCAATAACACGAGCAACAATCA 1391

451  ThrProPheGluLeuIleGlnGlnArgGluThrLysGluValAspSerLy 467
|||||
1392  ACTCCTTTTGAATTTGATACAGAGAGAGAAACCAAGAGGTGGACAGCAA 1441

467  sGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheValG 484
|||||
1442  GGAACCTTTCTTATTTTGAATCTCAACCACATGATCTTGTGTTGTAG 1491

484  luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
|||||
1492  AGATGACGGCTCAAAAAAGTAATGCATGTTTCTTCAGCAGAACTGAATTAT 1541

501  SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnValLy 517
|||||
1542  TCACTGCCATATGACTTAAACACCAAAATACGTAATGCCTCTAATGCTAAA 1591

517  sHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValG 534
|||||
1592  GCACCATGACTCTAGTGCTCTTGGTGATATATTCTTACATACCTTTAGTGG 1641

534  luAsnProTyrPheSerSerTrpProProSerGlyThrSerSerLysMet 550
|||||
1642  AAAATCCTTATTTTTCATCATGCTCCCAAGTGGTACCAGTCTTAAGATG 1691

551  SerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerLe 567
|||||
1692  TCTCTTGATTTTACCTGAGAAGAGATGGAAGTGTGTTTTTCTCTCTCTCT 1741
```


567 uLeuProThrSerSerThrSerLeuPheSerTyrTrpAsnSerHisSerS 584
 1742 GTTGCCAACTCCTACATCCTCTCTCTATTACATTCACATAGTT 1791
 584 eLeuSerLeuAsnSerProThrAsnLeSerSerLeuLeuAsnGlnGlu 600
 1792 CTTTATCACTGAATTCCTCAACCAATATTTCTCATTATTTGAACAGGAG 1841
 601 SerAlaValLeuAlaThrAlaProArgLeaAspAspGluLeuProPro 617
 1842 TCAGCTGTACTAGCAACTCTCTCAAGGATAGATGATGAATCCCTCC 1891
 617 oLeuProValArgThrProGluSerPheLeValValGluAlaGlyG 634
 1892 ACTTCTGTACGACACCTGATCATTTATCTGCTGAGGAAGCTGGAG 1941
 634 luPheSerProAsnValProLysSerLeuSerSerAlaValLysValLys 650
 1942 AATTTCTACCAATGTTCCCAATCCTTATCTCAGCTGTGAAGGTAAA 1991
 651 lLeGlyThrSerLeuGluTrpGlyGlyThrSerGluProLysLysPheAs 667
 1992 ATTGAACATCACTGGAATGGGTGGACATCTGAACCAAGAAATTTGA 2041
 667 pAspSerValLLeuLeuArgProSerLysSerValLysLeuArgSerProL 684
 2042 TGACTCTGTGATACITAGACCAAGCAAGAGTGTAAACTCCGAATCCTA 2091
 684 ysSerGlyLysAsnPheSerTrpLeu 692
 2092 AATCAGGTAAAAATTTCTCTTGGCTT 2117

seq_name: /net/abss06/STDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:AAx90695

seq_documentation_block:

ID AAx90695 standard; cDNA; 3058 BP.
 AC AAx90695;
 DT 01-OCT-1999 (first entry)
 DE Human Cytoplasmic phosphatase Lyp 1, cDNA.
 KW Lymphoid protein Tyrosine Phosphatase gene; Lyp gene; lymphoid cell;
 KW intracellular tyrosine phosphatase; PTPase; lymphocyte; thymocyte;
 KW T cell; B cell; protein tyrosine kinase; PKs; immunosuppressant;
 KW lymphoma; transgenic animal; T cell antigen receptor signalling;
 KW cytokine receptor signalling; protein tyrosine phosphatase; PTPase; ds.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 42..2468
 FT /*tag= a
 FT /product= "Cytoplasmic phosphatase, Lyp 1"
 FT /note= "Expressed in thymocytes, mature T and B cells"
 XX
 PN W09936548-A1.
 XX
 XX 22-JUL-1999.
 XX
 XX 18-JAN-1999; 99WO-CA00038.
 XX
 XX 16-JAN-1998; 98CA-2220853.
 XX
 XX (HSC-) HSC RES & DEV LP.
 XX
 XX Roifman CM;
 XX
 XX WPI; 1999-444404/37.
 DR P-PSDB; AAY28652.
 XX

PT New nucleic acid encoding intracellular tyrosine phosphatase and
 PT related proteins, used to modulate signalling through T cells,
 PT particularly as immunosuppressant

XX Claim 5a; Page 52; 105pp; English.

XX The present cDNA sequence is that of the Lymphoid Protein Tyrosine
 CC Phosphatase gene, Lyp 1 that encodes a cytoplasmic tyrosine phosphatase
 CC protein with a single catalytic domain. The Lyp gene has been localised
 CC to human chromosome 1p13. It is expressed in the lymphoid cells,
 CC particularly in thymocytes, mature B and T cells. It is involved in
 CC lymphocyte growth and development. Lyp gene products are important for
 CC regulation of T cell antigen and cytokine receptor signalling and early
 CC and late stages of T cell differentiation. Lyp 1 has immunosuppressive
 CC activity. This sequence is used to treat lymphoma, to reduce or prevent
 CC T cell activation or proliferation and to control thymocyte
 CC differentiation. Fragments of this gene are used, as probes or primers,
 CC to identify allelic variations and for diagnosis of diseases associated
 CC with reduced Lyp activity. Transgenic animals in which a Lyp gene has
 CC been inactivated or replaced by a heterologous Lyp gene are used to
 CC study and screen for possible therapeutic agents.

XX Sequence 3058 BP; 1015 A; 592 C; 553 G; 898 T; 0 other;

alignment_scores:

Quality: 3573.00 Length: 685
 Ratio: 5.216 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-600-358A-4 x AAx90695 ..

Align seg 1/1 to: AAx90695 from: 1 to: 3058

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLys 17
 42 ATGGACCAAGAGAAATTCGCAGAAATTCCTGGATGAGGCCCAAGCAA 91
 17 sLysIleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArgG 34
 92 GAAATATTACTAAAGAGGAGTTTGGCAATGAATTTCTGAAGCTGAAAGGC 141
 34 InSerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50
 142 ATCTACCAAGTACAGGACAGCAAAACCTATCCTACAACTGTGCTGAG 191
 51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
 192 AATGCCAAGAATATCAAGAAAAACAGATATAAGGATATTTGCCCTATGA 241
 67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSert 84
 242 TTATAGCCGGGTAGAACTATCCCTGATAACCTCTGATGAGGATTCAGCT 291
 84 YrIleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
 292 ACATCAATGCCAATTCATTATAGGAGTTTATGGACCCCAAGCTTATATT 341
 101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetIl 117
 342 GCCACCCAGGGTCCCTTTATCTACAAACCTCTCTGGACTTCGTGAGGATGAT 391
 117 eTrpGluTyrSerValLeuIleIleValMetAlaCysMetGluTyrGluM 134
 392 TTGGGAATATAGTGTCTTATCAATTGTTATGGCATGCGATGGAGTATGAA 441
 134 etGlyLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
 442 TGGGAAAGAAAAAGTGTGAGCGCTACTGGGCTGAGCCAGGAGATGCGAG 491
 151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
 |||||

492 CTGGAATTTGGCCCTTCTCTGTATCTCTGAAAGCTGAAAAAGGAATC 541
167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
|||||
542 TGATTATATATCAGGACTCTAAAGTTAAGTCAATAGTGAACCTGAA 591
184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSer 200
|||||
592 CTATCTACCAAGTTTCATTACAAGATTGGCCAGACCATGATGACCTTCA 641
201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnG1 217
|||||
642 TCTATAGACCTTATCTTGGAGCTCATCTGGGATGTCAGTTGTTACCAAGA 691
217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
|||||
692 GGATGACAGTGTCCCATATGCTATCTACTGCAGTGTGGCTGGGAAGGA 741
234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
|||||
742 CTGGTGTATTGCTGCTATTGTTGATTATACATGGATGTTGCTAAAGAT 791
251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267
|||||
792 GGGATAATTCCTGAGAACTTCAGTGTGTTTCAGTTGATCGGGAATGCG 841
267 qThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValT 284
|||||
842 GACACAGAGGCTTCATTAGTTCAAAGCAGGAACAATATGAACCTGGTCT 891
284 YrAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
|||||
892 ACATGCTGTATTAGAACTATTAAAGACAGATGGATGTTATCAGAGAT 941
301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
|||||
942 AACATTTCTGGAACAGAGAGTCAAGCAAGACATTTGTTCTCTCGAGAAAA 991
317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
|||||
992 TCACACTCTCCAGCAGACTCTTATCTCTCTTAATTTACCAAAAGTACCA 1041
334 hrLysAlaAlaLysMetMetAsnGlnGlnArgThrLysMetGluIleLys 350
|||||
1042 CAAAAGCAGCAAAATGATGAACCAACAAAGGACAAATGGAATCAAA 1091
351 GluSerSerSerPheAspPheArgThrSerGluIleSerAlaLysGluG1 367
|||||
1092 GAATCTTCTCTCTTGTGACTTTAGGACTTCTGAATAAGTGCAAAAGAGA 1141
367 uLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGluL 384
|||||
1142 GCTAGTTTGGCCCTGCTTAATCAACACTCTCTTTGACTTCTGGAGC 1191
384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnThr 400
|||||
1192 TAATATACAGTTTGCACAAATCTCTGACACCAACCATGAATGGCAGACA 1241
401 LysAlaPheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAs 417
|||||
1242 AAGGCATTTCCAAATAGTTGGGAGCCTCTCTCAGAAGCATCAAAAGTTTGA 1291
417 pLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsnA 434
|||||
1292 TTTGGGCTCTCTTTGTTGGAGGATGTTCTAATCTAAACCTGTAATG 1341
434 laAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSer 450
|||||
1342 CAGCAGGAAGATATTTAATTCAAAGTGCCAAATACACGGACCAATCA 1391
451 ThrProPheGluLeuIleGlnArgGluThrLysGluValAspSerLy 467
|||||
1392 ACTCCTTTTGAATTTGATACAGAGAGAAACCAAGAGGTGGACAGCAA 1441

467 sGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheValG 484
|||||
1442 GGAACACTTTTCTATTTGGAATCTCAACCATGATCTCTGTTTGTAG 1491
484 luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
|||||
1492 AGATGAGGCTCAAAAAAGTAATGCATGTTTCTTTCAGCAGAACTGAATTAT 1541
501 SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnValLy 517
|||||
1542 TCCTGCTCATATGACTCTAAACACCAAAATACGTAATGCCTCTAATGTA 1591
517 sHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValG 534
|||||
1592 GCACCATGACTCTAGTCTTGGTGTATATCTTACATACCTTTAGTGG 1641
534 luAsnProTyrPheSerSerTrpProSerGlyThrSerSerLysMet 550
|||||
1642 AAAATCCTTATTTTTCATCATGCTGCTCAAGTGTACCACTTCTTCTCT 1691
551 SerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerLe 567
|||||
1692 TCTCTTGATTTACCTGAGAAAGCAAGATGGAACCTGTTTCTCTTCTCT 1741
567 uLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSerS 584
|||||
1742 GTTGCCCAACATCTCTACATCCCTCTCTCTTATTACAAATCACAATG 1791
584 erLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnGlu 600
|||||
1792 CTTTATCAGTGAATTTCTCAACCAATATTTCTCCTACATTTGAACCAAG 1841
601 SerAlaValLeuAlaThrAlaProArgIleAspAspGluIleProProPr 617
|||||
1842 TCAGCTGTACTAGCACTGCTCAAGGATAGATGAATGAAATCCCCCTCC 1891
617 oLeuProValArgThrProGluSerPheIleValValGluAlaGlyG 634
|||||
1892 ACTTCTGTACGACACCTCAATCATTTATTTGTTGAGGAGCTGGAG 1941
634 luPheSerProAsnValProLysSerLeuSerSerAlaValLysValLys 650
|||||
1942 AATTCTCACCAAATGTTCCCAAAATCTTATCTCAGCTGTGAAGGTAAA 1991
651 IleGlyThrSerLeuGluTrpGlyGlyThrSerGluProLysLysPheAs 667
|||||
1992 ATTGGAACATCAGTGGAAATGGGTGGAAACATCTGAACCAAGAAATTTGA 2041
667 pAspSerValIleLeuArgProSerLysSerValLysLeuArgSerProL 684
|||||
2042 TGACTCTGTACTACTAGACCAAGAGTGTAAACTCCGAAGTCTTA 2091
684 ysSer 685
|||||
2092 AATCA 2096

seq_name: /net/abss06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT:AAV81742

seq_documentation_block:

ID AAV81742 standard; cDNA; 3580 BP.

XX AAV81742;

AC

XX

DT 10-MAR-1999 (first entry)

XX

XX Human PTP04 encoding cDNA.

XX

XX

KW PTP04; PTP05; PTP10; SAD; ALP; ALK-7; protein tyrosine phosphatase;

KW type I receptor serine/threonine kinase; cancer; leukemia; lymphoma;

KW neurodegenerative disease; neuronal survival; Alzheimer's disease;

KW Parkinson's disease; Huntington's disease; ss.

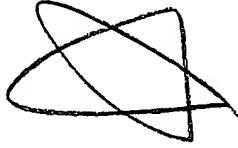
XX

OS Homo sapiens.
 XX Key Location/Qualifiers
 FH 53..2476
 FT CDS /*tag= a
 XX
 PN MO9849317-A2.
 XX
 XX 05-NOV-1998.
 XX
 XX 27-APR-1998; 98WO-US08439.
 XX
 XX 23-OCT-1997; 97US-0063595.
 PR 28-APR-1997; 97US-0044428.
 PR 20-MAY-1997; 97US-0047222.
 PR 11-JUN-1997; 97US-0049477.
 PR 11-JUN-1997; 97US-0049756.
 PR 18-JUN-1997; 97US-0049914.
 XX
 XX (SUG-) SUGEN INC.
 XX
 XX App H, Clary D, Courtneidge SA, Hui TH, Jallal B;
 PI Markby D, Onrust S, Peles E, Plowman GD;
 XX
 DR WPI; 1999-009434/01.
 DR P-PSDB; AAW89247.
 XX
 XX New nucleic acid encoding specific protein tyrosine phosphatases -
 PT useful for identifying specific modulators for treatment and
 PT prevention of cancer and neurodegenerative disease
 XX
 XX Claim 2; Page 145-146; 193pp; English.
 XX
 XX The present invention describes isolated, enriched or purified nucleic
 CC acids encoding PTP04, SAB, PTP05, PTP10, ALP and ALK-7 proteins. The
 CC present sequence encodes human PTP04. The above proteins, other than
 CC ALK-7, are protein tyrosine phosphatases (PTPs) and are used to identify
 CC substances that modulate their activity (i.e. agonists and antagonists,
 CC including NBP) in vivo or in vitro. These substances are used to treat
 CC or prevent diseases associated with abnormal signal transduction
 CC pathways that involve the proteins, particularly cancer (e.g. leukaemia
 CC and lymphoma), while modulators of ALK-7 (which is a type I receptor
 CC serine/threonine kinase) are used to promote neuronal survival,
 CC particularly for treating Alzheimer's, Parkinson's or Huntington's
 CC diseases. Nucleic acid fragments of the polynucleotides encoding the
 CC proteins can be used as probes to identify and clone related sequences;
 CC to detect protein-encoded RNA; to generate transgenic animals and in
 CC gene therapy (optionally after mutation). Ab are used to determine the
 CC proteins.
 XX
 SQ Sequence 3580 BP; 1213 A; 666 C; 618 G; 1083 T; 0 other;

alignment_scores:
 Quality: 3543.50 Length: 685
 Ratio: 5.188 Gaps: 1
 Percent Similarity: 99.708 Percent Identity: 99.416

alignment_block:
 US-09-600-358A-4 x AAV81742 ..
 Align seg 1/1 to: AAV81742 from: 1 to: 3580

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17
 53 ATGGACCAAGAGAGAAATTCGCAAGATTCCTGGATGAGGCCCAAGCAA 102
 17 sLysIleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArg 34
 103 GAAATTTACTTAAGAGAGGAGTTTGCCTCAATGATTTCTGAAGCTGAAAGGC 152
 34 InSerThrLysTyrLysAlaAspLysThrTyrProThrValAlaGlu 50
 |||||||||||||||||||||||||||||||||||||||||||||||||||



153 AATCTACCAAGTCAAGGCAGACAAAAACCTATCTACAACCTGTGGCTGAG 202
 51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
 : : : |||||||||||||||||||||||||||||||||||||||||||
 203 AGCCCAAGAAATATCAAGAAAAACAGATATTAAGGATATTTGGCCCTATGA 252
 67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSert 84
 |||||||||||||||||||||||||||||||||||||||||||
 253 TTATAGCCGGGTAGAACTATCCCTGATAACCTCTGATGAGGATCCAGCT 302
 84 yTlleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
 |||||||||||||||||||||||||||||||||||||||||||
 303 ACATCAATGCCAACTTCATTAAAGGAGTTTATGACCAAGGCTTATATT 352
 101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetI 117
 |||||||||||||||||||||||||||||||||||||||||||
 353 GCCACCCAGGGTCTCTATCTACAAACCTCTCTGGACTTCTGGAGGATGAT 402
 117 eTrpGluTyrSerValLeuIleIleValMetAlaCysMetGluTyrGlu 134
 |||||||||||||||||||||||||||||||||||||||||||
 403 TTGGGAATATAGTCTCTTATCATTTATGGCATGATGGAGTATGAAA 452
 134 etGlyLysLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
 |||||||||||||||||||||||||||||||||||||||||||
 453 TGGGAAAAGAAAAGTGTGAGCGCTACTGGGCTGAGCCAGGAGAGATGCG 502
 151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
 |||||||||||||||||||||||||||||||||||||||||||
 503 CTGGAAATTTGGCCCTTCTCTGTATCTGTGAAGCTGAAAAAGGAAATC 552
 167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
 |||||||||||||||||||||||||||||||||||||||||||
 553 TGATTATATATAATCAGGACTCTAAAAGTTAAGTTCAATAGTGAACCTCAA 602
 184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSer 200
 |||||||||||||||||||||||||||||||||||||||||||
 603 CTATCTACCAGTTTCATTACAAGAAATTTGGCAGACCATGATGTACCTTCA 652
 201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnG 217
 |||||||||||||||||||||||||||||||||||||||||||
 653 TCTATAGACCCCTATTCCTGTGAGCTCATCTGGGATGATGCTTTGTACCAAGA 702
 217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
 |||||||||||||||||||||||||||||||||||||||||||
 703 GGATGACAGGTGTTCCCATATGCATTCACCTGCAGCTGCTGGCTGGAAGGA 752
 234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
 |||||||||||||||||||||||||||||||||||||||||||
 753 CTGGTGTATTGTGCTATT...GATTATACATGGATGTTGCTAAAAGAT 799
 251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267
 |||||||||||||||||||||||||||||||||||||||||||
 800 GGGATAATTCCTGAGAACTTCAGTGTTCAGTTTGATCCGGGAAATGCG 849
 267 qThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValT 284
 |||||||||||||||||||||||||||||||||||||||||||
 850 GACACAGAGGCTTCATTAGTTCAACAGCAGGAAACAATATGAACCTGGTCT 899
 284 yTrsAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
 |||||||||||||||||||||||||||||||||||||||||||
 900 ACAATGCTGTATTAGAACTATTTAAGACAGAGATGGATGTTATCAGAGAT 949
 301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
 |||||||||||||||||||||||||||||||||||||||||||
 950 AAACATTCTGGACAGAGAGTCAAGCAAGCAATGTTATCTCTGAGAAAAA 999
 317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
 |||||||||||||||||||||||||||||||||||||||||||
 1000 TCACACTCTCCAAGCAGACTCTTATTTCTCTCTCTCTCTCTCTCTCTCT 1049
 334 hrLysAlaIleLysMetMetAsnGlnGlnArgThrLysMetGluIleLys 350
 |||||||||||||||||||||||||||||||||||||||||||
 1050 CAAAAGCAGCAAAAATGATGAACCAACAAAGGACAAAATGGAATCAAA 1099

```
351 GluSerSerSerPheAspPheArgThrSerGluIleSerAlaLysGluG1 367
|||||
1100 GAATCTCTCTCTTGCACCTTACCTTCTGAAATAGTGCAAAAGAGA 1149
|||||
367 uLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGluL 384
|||||
1150 GCTAGTTTTCACCCCTGCTAAATCAAGCACTTCTTTGACTTTCTGGAGC 1199
|||||
384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnThr 400
|||||
1200 TAAATACAGTTTGTGACAAATGCTGCACACCAACCATGAATGGCAGACA 1249
|||||
401 LysAlaPheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAs 417
|||||
1250 AAGGCATTTCCAAATAGTTGGGAGCCTCTTCAGACGATCAAAAGTTTGA 1299
|||||
417 pLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsnA 434
|||||
1300 TTTGGGCTCTCTTTTGTGAGGATGTTCTTAATTTCTAAACCTGTAATG 1349
|||||
434 laalaglyArgTyrPheAsnSerLysValProIleThrArgThrLysSer 450
|||||
1350 CACGAGAAGATATTTAAATTCAAAGGTGCCAATAACACGGACCAATCA 1399
|||||
451 ThrProPheGluLeuIleGlnGlnArgGluThrLysGluValAspSerLy 467
|||||
1400 ACTCCTTTGAAATGTATACAGCAGAGAGAAACCAAGAGGTGGACAGCA 1449
|||||
467 sGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheValG 484
|||||
1450 GGAATACTTTCTTATTTGGAATCTCAACACCATGATTTCTGTTTGTAG 1499
|||||
484 luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
|||||
1500 AGATGAGGCTCAAAAAGTAAATGCATGTTCTTCAGCAGCAACTGAATTA 1549
|||||
501 SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnVally 517
|||||
1550 TCAGTCCATATGACTTAACACCAAAATACGTAATGCCCTCTAATGTAA 1599
|||||
517 sHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValG 534
|||||
1600 GCACCATGACTCTAGTCTCTTGGTGATATTTCTTACATACCTTTAGTGG 1649
|||||
534 luAsnProTyrPheSerSerTrpProSerGlyThrSerSerLysMet 550
|||||
1650 AAATCCTTATTTTTCATCATGCGCTCCCAAGTGGTACCAGTTCTAAGATG 1699
|||||
551 SerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerLe 567
|||||
1700 TCTCTTGATTTTACCTGAGAGCAAGATGGAACTGTTTTCCTTCTCTCT 1749
|||||
567 uLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSerS 584
|||||
1750 GTTGCCAAATCCTCTACATCCCTCTCTCTTCTTATTAACAATTCACATG 1799
|||||
584 erLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnGlu 600
|||||
1800 CTTTATCAGTGAATTTCTCAACCAATATTTCTCACTATTATGAACCAAGG 1849
|||||
601 SerAlaValLeuAlaThrAlaProArgIleAspAspGluIleProProPr 617
|||||
1850 TCAGCTGTACTAGCACTGCTCCAGGATAGATGAAATCCCCCTCC 1899
|||||
617 oLeuProValArgThrProGluSerPheIleValValGluGluAlaGlyG 634
|||||
1900 ACTTCTGTACGGACACCTCAATCATTTATTGTGGTTGAGGAAGCTGGAG 1949
|||||
634 luPheSerProAsnValProLysSerLeuSerSerAlaValLysValLys 650
|||||
1950 AATTCTCACCAAAATGTCCCAAAATCTTATCTCAGCTGTGAAGGTAATA 1999
|||||
```

```
651 ileGlyThrSerLeuGluTrpGlyGlyThrSerGluProLysLysPheAs 667
|||||
2000 ATTGGAACATCACTGGAATGGGTGGAACATCTGAACCAAGAAATTTGA 2049
|||||
667 pAspSerValIleLeuArgProSerLysSerValLysLeuArgSerProL 684
|||||
2050 TGACTCTGTGATACTTAGACCAAGCAAGAGTGTAAAACTCCGAAGTCCTA 2099
|||||
684 ysSer 685
|||||
2100 AATCA 2104
|||||

seq_name: /net/abss06/sids1/gcgdata/hold-geneseg/geneseqn-emb1/NA2001A.DAT:AAH79342
seq_documentation_block:
ID AAH79342 standard; DNA; 2452 BP.
XX
AC AAH79342;
XX
DT 04-DEC-2001 (first entry)
XX
DE Human tyrosine phosphatase coding sequence.
XX
KW Human; tyrosine phosphatase; hPTP; ds.
XX
OS Homo sapiens.
XX
PN CN1302899-A.
XX
PD 11-JUL-2001.
XX
PF 29-OCT-1999; 99CN-0119935.
XX
PR 29-OCT-1999; 99CN-0119935.
XX
PA (NANF-) NANFANG RES CENT STATE HUMAN GENE GROUP.
XX
PI Ren S, Wu T, Qian B;
XX
DR WPI; 2001-550591/62.
DR P-PSDB; AAG78623.
XX
Human protein tyrosine phosphatase and its coding sequence -
PS Claim 1; Page 19-21(Disclosure); 29pp; Chinese.
XX
The present invention provides the protein and coding sequences of human
CC tyrosine phosphatase hPTP. The protein is expressed in human normal
CC suprarenal tissue. The present sequence is the coding sequence of the
CC invention.
XX
SQ Sequence 2452 BP; 791 A; 516 C; 468 G; 677 T; 0 other;

alignment_scores:
Quality: 2977.50 Length: 687
Ratio: 4.660 Gaps: 11
Percent Similarity: 93.013 Percent Identity: 89.229

alignment_block:
US-09-600-358A-4 x AAH79342 ..
Align seg 1/1 to: AAH79342 from: 1 to: 2452

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17
|||||
10 ATGGACCAAGAGAAATTCCTGCAGAAAGTTCTCGATGAGGCCCAAGCAA 59
|||||
17 sLysIleThrLysGluGluPheAlaAsnGluPheLeuLysLeuArg 34
|||||
60 GAAATTAATAAGAGAGGATTTGCCAATGAATTCGAACTGAAAGGC 109
|||||
34 InSerThrLysTyrLysAlaAspLysThrTyrProThrValAlaGlu 50
|||||
```

|||||
110 AATCTACCAAGTACAAAGCAGACAAAACCTATCTACAACTGTGGCTGAG 159
51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
:::
160 AAGCCCAAGAAATATCAAGAAAACAGATATAAGGATATTTGGCCCTATGA 209
67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSert 84
|||||
210 TTATAGCGGGTAGAACTATCCCTGATAAACCTCTGATGAGGATCCAGCT 259
84 yrlleAsnAlaAspPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
|||||
260 ACATCAATGCCAACTTCATTAAAGGAGGTTTATGACCCGAGGCTTATATT 309
101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheThrPArgMetI1 117
310 GCCACCCAGGGCTCTTATCTACAAACCTCTGGACTTCTGGAGGATGAT 359
117 eTrpGluTyrSerValLeuIleIleValMetAlaCysMetGluTyrGluM 134
|||||
360 TTGGGAATATAGTGCCTTATCATTTATGGCATGATGGAGTATGAAA 409
134 etGlyLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
|||||
410 TGGGAAAGAAAAGTGTGAGCGCTACTGGGCTGAGCCAGGAGATGCGAG 459
151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
460 CTGGAAATTGGCCCTTTCTCTGTATCTGTGAAGCTGAAAAGGAAATC 509
167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
|||||
510 TGATTATATATCAGGACTCTAAAAGTTAAGTTCAATAGTGAACCTCGAA 559
184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSer 200
560 CTATCTACCACTTTCATTACAAAGATTGGCCAGACCATGATGATACCTTCA 609
201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnG1 217
610 TCTATAGACCCATCTTGTAGCTCATCTGGGATGTCCGTTGATCCCAAGA 659
217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
660 GGATGACAGTGTCCCATATGCAATTCAGTCAGTGTGGCTGTGGAAGA 709
234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
710 CTGGTGTATTGTGCTATT...GATTATACATGGATGTGCTAAAGAT 756
251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267
757 GGGATAATTCTCGAGAACTTCAGTGTGTTTTCAGTTGTATCGGGAAATGCG 806
267 qThrGlnArgProSerLeuValGlnThrGlnGlnGlnTyrGluLeuValT 284
807 GACACAGGGCTTCATTAGTTCAAACGACGAGAACATATGAACCTGGTCT 856
284 yrlleAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
857 ACAATGCTGTATTAGAACTATTAAAGACAGACATGATGATTTATCAGAGAT 906
301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
907 AAACATTCTGGAACAGAGAGCTCAAGCAAGCATTTGATTCCTGAGAAAA 956
317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
957 TCACACTCTTCAACGACACTCTTATCTCTTAATTTTACCAAAAAGTACCA 1006
334 hrLysAlaAlaLysMetMetAsnGlnGlnArgThrLysMetGluIleLys 350
|||||

1007 CAAAGCAGCAAAAAATGATGAACCAACAAAGGACAAAAATGAAATCAAA 1056
351 GluSerSerPheAspPheArgThrSerGluIleSerAlaLysGluG1 367
|||||
1057 GAATCTCTCTCTTGTGACTTTAGGACTTCTGAAATAAGTGCAAAAGAGA 1106
367 uLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGluL 384
1107 CGTAAGTTTGGCCCG...GAAAACACGGCTTTTTCCTTCCTTCGGAAG 1153
384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnThr 400
1154 TAAATTCAGTTTGCAAAATGTTG...CCCCACCATGAAATGCCAGCA 1200
401 LysAlaPheProIleValGlyGluPro LeuGlnLysHisGlnSerLeuA 417
1201 AAGCTTTCCAAATAGTGGGA...AGCTCTTCAGAAGCATCAAAAT...T 1244
417 spLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValasn 433
1245 GAATGGGCCTCTCTTGTGTTAGGAATGTCTA...ATTCTAACCTTGAT 1291
434 AlaAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSe 450
1292 GCAGCAGAAAGATATTTAATCCAAAGTGCCAAT...ACCCGACCAATC 1338
450 rThrProPheGluLeuIleGlnArgGluThrLysGluValAspSerL 467
1339 AACTCCT.TTTGAATTGATCCAGCAA...GAGAACCAAGGAGGTGACAGCA 1384
467 ysGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheVal 483
1385 AGAAACTTCTTATTT...GAATCTCAACACATGATCTGTGTTGTA 1431
484 GluMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTy 500
1432 GAGATGAGGCTCAAAAAGTAATCATGTTTCTTCAGCAGAACTGAATTA 1481
500 rSerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnValL 517
1482 TTCCTGCCATATGACTCTAAACCAACCAATACGTAATGCCCTCTAATGTAA 1531
517 ysHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuVal 533
1532 AGCACCATGACTCTAGTGTCTCTGGTGATATATTCTTACATACCTTTAGTG 1581
534 GluAsnProTyrPheSerSerTrpProSerGlyThrSerSerLysMe 550
1582 GAAATCTCTTATTTTTCATCATGCGCTCCAAGTGTACCAAGTTCTAAGAT 1631
550 tSerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerL 567
1632 GTCTCTTGATTTACTGAGAAGCAAGATGGAACGTGTTTTCCTCTCTCTC 1681
567 euLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSer 583
1682 TGTGTCACACATCTCTACATCCCTCTCTCTTATTACAAATTCACATGAT 1731
584 SerLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnG1 600
1732 TCTTTATCACTGAATTTCTCCAACCAATATTTCTCAGTGTATGAACAGA 1781
600 uSerAlaValIleAlaThrAlaProArgIleAspAspGluIleProProp 617
1782 GTCAGCTGTACTAGCAACTGCTCCAAGGATAGATGATGAAAT.CCCCTC 1830
617 roLeuProValArg.ThrProGluSerPheIleValValGluGluAlaG1 633
1831 CACTTCTGTACGGGACACCTCAATCATTTATTGCTGTTGAGGAAGCTGG 1880
633 yGluPheSerProAsnValProLysSerLeuSerSerAlaValLysValL 650
1881 AGAATCTCACCAAAATGTTCCCAAAATCCTTATCTCAGTGTGAAGGTAA 1930
|||||

650 ysileGlyThrSerLeuGluTrpGlyGlyThrSerGluProLysLysPhe 666
|||||
1931 AAATTGGACATCACTGGAATGGGTGGACATCTGAACCAAGAAATTT 1980

667 AspAspSerValIleLeuArgProSerLysSerValLysLeuArgSerPr 683
|||||
1981 GATGACTCTGTGATCTTAGACCAAGCAGAGTGTAAAACTCGGAAGTCC 2030

683 oLysSer 685
|||||
2031 TAAATCA 2037

seq_name: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT:ABI99910

seq_documentation_block:

ID ABI99910 standard; cDNA; 2983 BP.

XX

AC

XX

DT

XX

DE

XX

KW

KW

XX

OS

XX

PN

XX

PD

XX

PF

XX

PR

XX

PA

PI

XX

DR

XX

PT

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Sequence 2983 BP; 923 A; 634 C; 681 G; 745 T; 0 other;

alignment_scores:

Quality: 1104.50

Ratio: 2.493

Percent Similarity: 59.623

Length: 743

Gaps: 22

Percent Identity: 35.935

alignment_block:
US-09-600-358A-4 x ABI99910 ..

Align seg 1/1 to: ABI99910 from: 1 to: 2983

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17
|||||
52 ATGGACCAAGTGGAGATCCTGAGGAGGTTCAATCCAGAGGTCAGGCCAT 101

17 slYsIleThr.....LysGluGluPheAlaAsnGluPheLeuL 30
|||||
102 GAAGAGTCGGGATCACAATGGGAGGAGCAACTTCGCCCGGGACTTCATGC 151

30 ysLeuLysArgGlnSerThrLysLysAlaAspLysThrTyrProThr 46
:||||:|||||
152 GATTGAGAAGATTGCTACCAATATAGAACAGAAAGATTTATATCCACA 201

47 ThrValAlaGluAsnAlaLysAsnIleLysLysAsnArgTyrLysAsp 63
:||||:|||||
202 GCCACTGGAGAAAAAGAAATGTTAAAGAACAGACAGATATAAGGACAT 251

63 eleuProTyrAspTyrSerArgValGluLeuSerLeuIleThrSerAsp 80
:||||:|||||
252 ACTGCCATTTGATCACAGCGGAGTTAAGTTGACTTTGAAGACTCCATCCC 301

80 luAspSerSerTyrIleAsnAlaAsnPheIleLysGlyValTyrGlyPro 96
:||||:|||||
302 AAGATTCAGATTATATCAATGCAAAATTTATTAAAGGTTGTTATGGGCCA 351

97 LysAlaTyrIleAlaThrGlnGlyProLeuSerThrThrLeuLeuAspPh 113
|||||
352 AAAGCATATGTGGCAACCCAGGGCCTTCCGGAATACAGTCATAGACTT 401

113 eTrpArgMetIleTrpGluTyrSerValLeuIleIleValMetAlaCysM 130
|||||
402 CTGGAGGATGATATGGAGTATAATGTTGTGATGATCGTATGGCGCTGC 451

130 etGluTyrGluMetGlyLysLysCysGluArgTyrTrpAlaGluPro 146
|||||
452 GAGAAATTTGAGATGGGAGGAAAGAGTGTAGCGCTACTGGCGCTTGTAT 501

147 GlyGluMetGlnLeuGluPheGlyProPheSerValSerCysGluAlaG 163
|||||
502 GGAGAAGATCCCTATAACATTTGCACCATTTAAATTTCTGTGAAATGA 551

163 uLysArgLysSerAspTyrIleIleArgThrLeuLysValLysPheAsnS 180
:||||:|||||
552 ACAAGCAAGAACCGACTACTTCTATCGGACACACTTTTACTGAATTTCAA 601

180 erGluThrArgThrIleTyrGlnPheHisTyrLysAsnTrpProAspHis 196
:||||:|||||
602 ATGAATCCCGTCGGCTCTATCAGTTTCAATTCATGACTGGCCAGGACCAT 651

197 AspValProSerSerIleAspSerValProIleLeuIleTrpAspValAr 213
| | | | |
652 GATGTTCTCTCGTCATTTGATTCTATCTCGACATGATAGCTTAATGAG 701

213 gCysTyrGlnGluAspAspSerValProIleCysIleHisCysSerAlaG 230
| | | | |
702 GAAATCCCAAGAACATGAAGATGTGCTATTTGTTATTCATTCGAGTGCTG 751

230 lyCysGlyArgThrGlyValIleCysAlaIleValAspTyrThrTrpMet 246
| | | | |
752 GCTGTGGACCAACAGGTCGTATTGTGGCCATA...GATTACAGCTGGAAC 798

247 LeuLeuLysAspGlyIleIleProGluAsnPheSerValPheSerLeuIl 263
| | | | |
799 TTACTGAAGACAGGAAATTCACAGAGGAATTTAATGATTTAATTAAT 848

263 eArgGluMetArgThrGlnArgProSerLeuValGlnThrGlnGluGlnT 280
:||||:|||||
849 ACAAGAAATGAGAACACAGAGGCACTCGGCAGTACAAACAAGAGCAGT 898

```
280 yrgluLeuValTyrAsnAlaValLeuGluLeuPheLysArgGlnMetAsp 296
|||||
899 ATGAACCTTGTCATAGGGCTATGCTCAACTGTTGAAACAGCTACAC 948
297 ValIle..... 298
949 TGTATGAATTCATGAGCACAGAAGATCCGTGATGGTAATGAAATTACC 998
299 .....ArgaspLysHisSer. 303
999 ACTGGAACATATGTCATGTCATTCATGATAGCGAGACGACCTCTCCGCCA 1048
304 .....GlyThr 305
1049 AAGCGCGCGGACTCGAAGTTGCCTTGTAGAAGGGGATGCCAAGGAAGA 1098
306 GluSerGlnAlaLysHisCysIleProGluLysAsnHisThrLeuGlnAl 322
1099 AATACTACAGCCACCAGAACCTCACCCGGTGCACCCCTCTGAGGCCAT 1148
322 aAspSerTyrSerProAsnLeuProLysSerThrThrLysAlaAlaLysM 339
1149 CACCTCCTTCACG.....CTTCCCAACGTTACCACTGTGTCAGGACAG 1192
339 et.....MetAsn 341
1193 TCAGAGGTACCAACCAAGCCAGTCTGCACATGGCCTCACCAGAGCAAC 1242
342 GlnGlnArgThrLysMetGluIleLysGluSerSerPheAspPheAr 358
1243 ACCCAGCGGACCTCAACAGAA.....GCTATGATAAATCAGCG 1280
358 qThrSerGluIleSerAlaLysGluLeuValLeuHisProAlaLysS 375
1281 GACC.....AATGGGCAAAAA 1297
375 erSerThrSerPheAspPheLeuGluLeuAsnTyrSerPheAspLysAsn 391
1298 GTGAATCAGCTATTGAGCACATA.....GATAAGAG 1329
392 AlaAspThrThrMetLysTrpGlnThrLysAlaPheProIle..... 405
1330 TTAGAGCGCAATTAAAGTTTGTAGATTAGAAAGTCCCTCTCCAAGAAGG 1379
405 ..... 405
1380 GCCCAAAAGTTTGTAGGGAACACACTCTTGAATAGGGGACATGCGATTA 1429
406 .....ValGlyGluProLeuGlnLysHisGln 414
1430 AAATTAAATCTGCTTCATCTCTGTTAGTTGACAGAACCTCTAAACCCACAG 1479
415 SerLeuAspLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysPr 431
1480 GAGTTAAAGTCAGAGTGGCCCTAAAGTTGATGATGATCTCAAGATTCTTG 1529
431 oValAsnAlaAlaGlyArgTyr..... 438
1530 CCGGACTGTAGTCGGCTCATCTACACAGAGCTGCTGAGTCGTGAGG 1579
439 .....PheAsnSerLysValProIleThrArgThrLysSerThrPro 452
1580 AGTCCCGAGCAACTCACACACCT...CCAGGCCACAGCTGCTTGCTT 1626
453 PheGluLeuIleGlnGlnArgGluThrLysGluValAspSerLysGluAs 469
1627 CTCGATAAG...AAAGGACAGTAACGTGGTCACTTCATGACCTGAAAA 1673
469 nPheSerTyrLeuGluSerGlnProHisAspSerCysPheValGluMetG 486
1674 TGCCACACCTGTACCCGACTCACCT...GACGGCAAAATCCCCAGATAATC 1720
```

```
486 InAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyrSerLeu 502
|||||
1721 ATTCACAGACTCTGAACCCGTGAGTTCCACACCC.....AACTCCACC 1764
503 ProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnValLysHisH 519
1765 GCAGAGAGGAGGAGCCAGCATCTT.....ACAGAGCACCA 1799
519 sAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValGluAsn 536
1800 CAACAGCTCC.....CCTCTGTTGAAAGCTC 1825
536 io.....TyrPheSerSerTyrProSerGly 545
1826 CCTCAGCTTTACCAACCTCTTCACTCTGAGGACTGGCACTCAGACGA 1875
546 ThrSerSerLysMetSerLeuAspLeuProGluLysGlnAspGlyThrVa 562
1876 GGGAGCTCT.....GATGGTGTCTGT 1895
562 lPheProSerSerLeuLeuProThrSerSerThrSerLeuPheSerTyrT 579
1896 GACCAAGGACAAACTAGCATTTCAACAGCAAGTCCACAGTGTCTCTCTG 1945
579 yrAsnSerHisSerSerLeuAsnSerProThrAsnIleSerSer 595
1946 CCAGTAGTCTCAGAGT...GCTTGCCATAGGAGAGATATGCCGATGCC 1992
596 LeuLeuAsnGlnGluSerAlaValLeuAlaThrAlaProArg..... 609
1993 ATTGCCAGACAGGAA.....GTAGCAGCAGCGGCATTCAGGTGC 2033
610 .....IleAspAspGluIleProProProLeuProVala 621
2034 TCAGAAAGATGCTGATGTTAGTAGGAGTGCCTCTCTCTTACCTGAAC 2083
621 rgThrProGluSerPheIleValValGlu 630
2084 GAACCTCTGAGCTCTTTGTTATAGCAGAT 2112
```

seq_name: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.AAF21801

seq_documentation_block:

ID AAF21801 standard; DNA; 1286 BP.

XX AAF21801;

XX 27-MAR-2001 (first entry)

DE Human breast and ovarian cancer associated antigen gene SEQ ID 188.

XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW neutropic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; anticancer; vulnery; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis; ds.
KW cardiovascular disorder; wound healing; neurological disease; ds.

XX Homo sapiens.

OS WQ200055173-A1.

PN 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05881.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

CC from a rat phaeochromocytoma PC12 cDNA library by PCR amplification
 CC using primers based on consensus sequences (see AAW49915-16) of known
 CC ptps. The invention relates to novel proteins (see AAW49906-14)
 CC involved in cellular signal transduction and to the nucleic acids
 CC (see AAV17097-99) coding for them, and provides vectors, host cells,
 CC purified recombinant proteins, methods for identifying compounds
 CC that bind the novel proteins or abrogate their interactions with
 CC natural binding partners, as well as methods for the diagnosis and
 CC treatment of diseases associated with the novel proteins. For
 CC ptp20, activators may act as anti-cancer therapeutics that
 CC stimulate cell differentiation rather than proliferation, while
 CC inhibitors may be useful for treating neural injuries by delaying
 CC the differentiation of transplanted neuronal stem cells until they
 CC are firmly grafted.

XX
 SQ Sequence 2226 BP; 574 A; 591 C; 604 G; 457 T; 0 other;

alignment_scores:
 Quality: 748.00 Length: 357
 Ratio: 3.041 Gaps: 7
 Percent Similarity: 68.908 Percent Identity: 43.137

alignment_block:

US-09-600-358A-4 x AAV17097 ..

Align seg 1/1 to: AAV17097 from: 1 to: 2226

```

3 GlnArgGluLeuLeuGlnLysPheLeuAspGluAlaGlnSerLysLysI 19
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
37 CAATCGGACCTAGTGAGGAGCTCTTGGAGCAGCAGGAGCGCGGACCA 86
19 eThrLysGluGlu...PheAlaAsnGluPheLeuLysLysLysArgGln 35
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
87 CGGGAAGGGGGCAATCCCTCGCCGGTGTTCAGCAGCATTAAGGCCCGCT 136
35 eThrLysLysLysAlaAspLysThrTyProThrValAlaGluAsn 51
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
137 CAGTGGCTTGAAGACTGAAGGTGTGCTCCACTTAAGCCGGCAGTCAG 186
52 AlaLysAsnIleLysLysAsnArgTyLysAspIleLeuProTyAspTy 68
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 CAGGGAACCTCAAGAAGAACCGCTACAAAGACGTGTACCGTATGATGA 236
68 rSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSeryI 85
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
237 GACGAGAGTCATCCCTTCCCTCCAGGAGGAGGACACGGAGATTACA 286
85 LeAsnAlaAsnPheIleLysGlyValTyGlyProLysAlaTyIleAla 101
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
287 TTAATGCCAACTTATCCGGGCACAGATGGNAGCCAGCCCTACATTCG 336
102 ThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetIleTr 118
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
337 AGCGAAGGACCCCTCCCTCACACTCTGTGGACTTCTGGCGCTGTGTTG 386
118 pGluTySerValLeuIleValMetAlaCysMetGluTyGluMetG 135
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
387 GGAGTTTGGAAATCAAGGTGATCTTGTGCTGTGCTGAGGAGCAGAAAATG 436
135 LyLysLysLysCysGluArgTyTrpAlaGluProGlyGluMetGlnLeu 151
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
437 GACGGAGGAAGTGTAGCTACTGGGCCCGCCAGGACGGGAG...CCTCTA 483
152 GluPheGlyProPheSerValSerCysGluAlaGluLysArg...LysSe 167
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
484 CAGGCGGGGCCCTTCTGCATCACCTCCACAAAGGAGACAGCAGCTGCTC 533
167 rAspTyIleLeuArgThrLeuLysValLysPheAsnSerGluThrArgT 184
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
534 GGACATCACTCTCAGGACCCCTCCAGGTTCATATTCAGAAGGAATCCGCTC 583
184 hrIleTyGlnPheHisTyLysAsnTrpProAspHisAspValProSer 200

```

```

584 CTGTGCACACAGCTACAGTACATGTCTTGGCGGACACACGGGTTCCACG 633
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyGlnG 217
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
634 AGTTCCGATCATATCTCACCATGGTGGAGGAGGCCGCTGCCTCCAAG 683
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArg 234
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
684 ACTTGGACCTGGACCCCTCTGTGCTCCACTGAGTGTGGCTGGCTGGAC 733
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
234 hrGlyValIleCysAlaIleValAspTyThrTrpMetLeuLeuLysAs 250
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
734 CAGGTGTCTTGTGTGCT...GTTGATTACGTGAGGAGTGTCTTCTGACT 780
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMet 267
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
781 CAGACAATCCCAACCAATTTGAGCTCTTTGAAGTGGTCTGGAGATGCG 830
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
267 gThrGlnArgProSerLeuValGlnThrGlnGluGlnTyGluLeuVal 284
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
831 GAACAGGACCTGCGCGGTGCACAGAGAGGAGGAGTACAGGTTCCTGT 880
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
284 yrAsnAlaValLeuLeuPheLysArgGlnMetAsp..... 296
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
881 ACCACACAGTGGCTCAGCTATTCTCCCGCAGCTCTCCAGAACACAGTCC 930
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
297 .....ValIleArgAspLysHis 302
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
931 CTCTACCAAGAACCTCAAGGAGAACCGCGCTCCAATCTGCAAGGATTC 980
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
302 sSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAsnHis 319
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
981 GTCCCTCAGGACCTCTCTCAGCC.....CTGCCTGCCACATCCCGCC 1021
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
319 hrLeuGlnAlaAspSerTySer.....ProAsnLeu 329
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1022 CACTGGGTGGCTTCTCAGGAGCATCTCGTGGCTGGCCACCGACCTT 1071
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
330 ProLysSerThrThrLysAla 336
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1072 CCCATGGCTGACACTTACGCT 1092
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

seq_name: /net/abss06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT:AAV03112

seq_documentation_block:

ID AAV03112 standard; cDNA; 1529 BP.

XX

AC AAV03112;

XX

DT 09-APR-1998 (first entry)

XX

XX cDNA encoding a novel non-receptor tyrosine phosphatase.

DE

XX Non-receptor protein tyrosine phosphatase; hematopoietic stem cell;

KW PTP HSC; progenitor cell; tyrosine phosphatase domain; PTP HSC agonist;

KW tyrosine dephosphorylation; tyrosine phosphatase; PTP HSC antagonist;

XX stem cell differentiation; ss.

OS

XX Mus sp.

XX

PH Key Location/Qualifiers

FT CDS 19..1380

FT /*tag= a

XX

XX WO9735019-A1.

XX

XX 25-SEP-1997.

XX

XX 17-MAR-1997; 97WO-US05278.

XX

XX 22-MAR-1996; 96US-0620526.

XX


```

FT CDS /note= "GC-rich track, part of Kozak sequence"
FT 44..1420
FT /*tag= b
FT /transl_except= (pos:69..71, aa:Arg)
FT 122 GCAGCGAGTTTCAGGACATCCAGCGCTCTCGCGCGCTGGAGGCTGA 171
FT /transl_except= (pos:162..164, aa:His)
FT 41 pLysThrTyrProThrThrValAlaGluAsnAlaLysAsnLleLysLysA 58
FT 172 CGCGCTGTCTCCACCGTGGCGGCGAGTCGCCAGAGACGTGAGGAAGA 221
FT /*tag= c
FT /note= "site of 367 bp intron"
FT 58 snArgTyrLysAspLleLeuProTyrAspTyrSerArgValGluLeuSer 74
FT 222 ACCGCTACAAAGACGTGCTTATGATCAGCGCGAGTAATCTCTCC 271
FT /*tag= d
FT /note= "site of 80 bp intron"
FT 75 LeuLleThrSerAspGluAspSerSerTyrIleAsnAlaAsnPhellely 91
FT 272 CTGCTCCAGGAGAGGACACAGCGACTACATTATGCAACTTCATCCG 321
FT /*tag= e
FT /note= "site of 91 bp intron"
FT 91 sGlyValTyrGlyProLysAlaTyrIleAlaThrGlnGlyProLeuSerT 108
FT 322 GGGCGTGGATGGAAGCCTGGCTTACATTGCCACGCAAGGACCTTGCCTC 371
FT /*tag= f
FT /note= "f-rich sequence required for
FT 2740..2747
FT polyA_signal /tag= g
FT 108 hrThrLeuLeuAspPheTrpArgMetIleTrpGluTyrSerValLeuile 124
FT 372 ACACCTGCTAGACTTCTGGAGACTGGTCTGGGAGTTTGGGGTCAAGGTG 421
FT /*tag= 97WO-IB00946.
FT 125 IleValMetAlaCysMetGluTyrGluMetGlyLysLysLysCysGluar 141
FT 422 ATCTCGATGGCTGTGAGAGATAGAGAAATGGCGGAAAAGGTGTAGCGG 471
FT 141 gTyrTrpAlaGluProGlyGluMetGlnLeuGluPheGlyPropheServ 158
FT 472 GTACTGGGCCAGGAGGAGGAG...CCACTGCAGAGTGGGCTTTTCTGCA 518
FT 158 alSerCysGluAlaGluLys...ArgLysSerAspTyrIleIleIlearghr 173
FT 519 TCACCTCTGATAAAGAGAGAGTGGCTGAATGAGGACATCATGCTCAGGACC 568
FT 174 LeuLysValLysPheAsnSerGluThrArgThrIleTyrGlnPheHisty 190
FT 569 CTCAGGTCACATTCAGAGAGGAGTCCCGTCTGTGTACCAGCTACAGTA 618
FT 190 rLysAsnTrpProAspHisAspValProSerSerIleAspProIleLeuG 207
FT 619 TATGCTTGGCCAGACCGTGGGTGCCAGCAGCTCTGACCATCATGCTCG 668
FT 207 LuLeuLleTrpAspValArgCysTyrGlnGluAspAspSerValProIle 223
FT 669 CCATGCTGGAGGAGCCGCTCCCTCCAGGGATCTGGCCCTGAACCCCTC 718
FT 224 CysIleHisCysSerAlaGlyCysGlyArgThrGlyValIleCysAlaI 240
FT 719 TGTGTCCACTGCAGTGGGGTGTGGGGAACAGCGGCTCTGTGTGCACC... 766
FT 240 eValAspTyrThrTrpMetLeuLysAspGlyIleIleProGluAsnP 257
FT 767 .GTGCATTATGTGAGCAGCTCTCTGACCCAGATGATCCACCTGACT 815
FT 257 heSerValPheSerLeuIleArgGluMetArgThrGlnArgProSerLeu 273
FT 816 TCAGTCTCTTGTATGTGCTTAAAGATCAGGAACGCGGCTCGGCC 865
FT 274 ValGlnThrGlnGluGlnTyrGluLeuValTyrAsnAlaValLeuLle 290
FT 866 GTGCAGACAGAGGAGCAGTACAGGTTCCTGTACCACACGGTGGCTCAGAT 915
FT 290 uPheLysArgGlnMetAspValIleArgAspLysHisSerGlyThrGluS 307
FT 916 GTTC.....TGCCTCCACACTCCAGAAATGCCAGCCCCCCTACCAGA 956
FT 307 exGlnAlaLysHisCysIlePro 314
FT 957 ACATCAAAAGAGAAATTGTGCCCA 979

```

alignment_scores: Quality: 708.00 Length: 291
 Ratio: 3.218 Gaps: 4
 Percent Similarity: 75.601 Percent Identity: 46.735

alignment_block:
 US-09-600-358A-4 x AAV17099 ..
 Align seg 1/1 to: AAV17099 from: 1 to: 2810

Claim 3; Fig 3a-d; 138pp; English.
 This cDNA sequence codes for a novel human protein tyrosine phosphatase (PTP), designated brain derived phosphatase 1 (BDP-1, see AAW49908), that is expressed in most tissues and cell lines at basal level, but expressed high in epithelium origin cell lines and cancer cell lines. BDP-1 was originally identified in a human brain cDNA library, although the full-length clone was isolated from the haematopoietic MEG01 cDNA library. The invention relates to novel proteins (see AAW49906-14) involved in cellular signal transduction and to the nucleic acids (see AAV17097-99) coding for them, and provides vectors, host cells, purified recombinant proteins, methods for identifying compounds that activate or inhibit the novel proteins, as well as methods for the diagnosis and treatment of diseases associated with the novel proteins.
 Sequence 2810 BP; 700 A; 808 C; 808 G; 494 T; 0 other;

(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 Aoki N, Chen Z, Kharitonov A, Kim YW, Nayler O;
 Ullrich A, Wang H;
 WPI: 1998-120302/11.
 P-PSDB: AAW49908.
 New phosphatase and kinase enzyme(s) - useful in the diagnosis and treatment of signal transduction disorders
 Claim 3; Fig 3a-d; 138pp; English.


```
308 nAlaLysHisCysIleProGluLysAsnHisThrLeuGlnAlaAspSert 325
    : : : : :
1611 GTACGGGAACATC.....ACCT 1627
325 ySerProAsnLeuProLysSerThrThrLysAlaAlaLys..... 338
    || || : : : : :
1628 ATCCCCCAGCATGAAGAAATCCCATGCCAAGGCTCCCGCACCTCGTCC 1677
339 .....MetMetAsnGlnGlnArgThrLy 346
1678 AAACACAGGAGGATGTGTATGAGAACTGCACACTAAGAACAGAGGGA 1727
346 smetGluIleLysGluSerSerPheAspPheArgThrSerGluIleS 363
    : : : : : || : : : : :
1728 GGAGAAAGTGAAGAGCAGCGTCCAGCAGACAGGAGAGCAAGGGTT 1777
363 erAlaLysGluGluLeuValLeuHisPro..... 372
    || || : : : : :
1778 CCCTCAAGAGGAGGAATTCGTGAGTCGAGGTACCGCGGCGCGGGAT 1827
373 .....AlaLysSerSerThrSerPheAsp..... 380
1828 CCACCGGTCCGCCACCATGTGTGAGCAAGGCGGAGGAGTGTACCGGGGT 1877
381 .....PheLeuGluLeuAsn.....TyrSerPheA 389
    : : : : : || : : : : :
1878 GGTGCCCATCTCTGTGTGAGCTGGACGGCGAGCTAAACGGCCACAAGTCA 1927
389 spLysAsnAlaAsp..... 393
    : : : : :
1928 GCGTGTCCGCGAGGCGGAGGCGATGCCACTACGGCAAGCTGACCCGTG 1977
393 ..... 393
1978 AAGTTCATCTCCACACCGGCAAGCTGCCGTGCCCTGGCCACCCCTCGT 2027
394 .ThrThrMetLysTrpGlnThrLysAlaPheProIleValGlyGluProL 410
    |||| : : : : :
2028 GACCACCTGACCTACGCGGTGAGTCTTACGCGGTACCCCGCACCA 2077
410 euGlnLysHisGlnSerLeuAspLeuGlySerLeuLeuPheGluGlyCys 426
    : : : : : || : : : : :
2078 TGAAGCAGCAGCACTTCTTCAAG.....TCCGCATGCCGGAAGGCTAC 2121
427 SerAsnSerLysProValAsnAlaAlaGlyArgTyrPheAsnSerLysVa 443
    : : : : :
2122 GTCGAGGAGCCACCATC.....TTCCTCAAGGACGACGG 2156
443 lProIleThrArgThrLysSerThrProPheGlu.....LeuIleG 457
    : : : : : |||| : : : : :
2157 CAACTACAAGACCGCGCCGCGAGGTGAAGTTCGAGGCGGACACCTGTGA 2206
457 lnGlnArgGluThrLysGluValAspSerLysGluAsnPheSerTyrLeu 473
    : : : : : || : : : : :
2207 ACCGATCTGAGCTGAAGGCGATCGACTTCAAGGAGGACGCGCAACATCCTG 2256
474 GluSerGln.....ProHisAspSerCysPheValGluMetGlnAl 487
    : : : : : |||| : : : : :
2257 GGGCAGACAGTGGAGTACAACTACAGACGCCACAACTCTATATCATGGC 2306
487 aGlnLysValMetHisValSerSerAlaGluLeuAsnTyrSerLeuProT 504
    : : : : : : : : : : :
2307 CGACAAG.....CAGAAGACGCGCATCAAGGTGAACCTTCAAGATC..... 2346
504 yrAspSerLysHisGlnIleArgAsnAlaSer..... 514
    : : : : : || : : : : :
2347 .....CGCCACACATCGAGGACGCGAGTTCAGCTCGCGGACAC 2388
515 .....AsnValLysHisHisAspSerSerAlaLeuGlyValTyrSe 528
    |||| : : : : :
2389 TACAGCAGAACACCCCATCGCGGCGGCGGCTGCTG..... 2427
528 rTyrIleProLeuValGluAsnProTyrPheSerSerTyrProProSerG 545
```

```
2428 .....CTGCCGACCACTACTGAGC..... 2451
545 lyThrSerSerLysMetSerLeuAspLeuProGluLysGlnAspGlyThr 561
    |||| : : : : :
2452 ..ACCAGTCGCGCTGAGCAAGACCCCAACGAGAGCGGATCCACATG 2499
562 Val 562
    |||
2500 GTC 2502
seq_name: /net/abss06/SIDSL/gcgdata/hold-geneseq/geneseqn-embl/NA1995.DAT.AAT06027
seq_documentation_block:
ID AAT06027 standard; cDNA; 5117 BP.
XX
AC AAT06027;
XX
DT 12-FEB-1996 (first entry)
XX
DE huDEP-1 cDNA.
XX
KW Density enhanced Type III receptor-like protein tyrosine phosphatase;
KW huDEP-1; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 350..4363
FT FT /*tag= a
XX
PN WO9530008-Al.
XX
PD 09-NOV-1995.
XX
PF 03-MAY-1995; 95WO-0505512.
XX
PR 03-MAY-1994; 94US-0237940.
XX
PA (COLD-) COLD SPRING HARBOR LAB.
XX
PI Oestman A, Tonks NK;
XX
DR WPI: 1995-393079/50.
DR P-PSDB; AAR85203.
XX
PT New density enhanced protein tyrosine phosphatase - used to develop
PT prods. to modify transcription, translation and/or activity of
PT tyrosine phosphatase(s).
XX
PS Claim 15; Page 27-34; 51pp; English.
XX
CC Degenerate primers (see AAT06028-29) corresp. to conserved protein
CC tyrosine phosphatase (PTP) amino acid sequences were used to prime
CC a PCR with a HeLa cell cDNA library as template. An isolated clone
CC was used to screen an oligo-dT-primed HeLa cDNA library. A cDNA
CC clone was obt'd. (given in AAT06027) that encoded a novel density-
CC enhanced Type III receptor-like PTP, designated huDEP-1 (AAR85203).
CC Expression of the cDNA in host cells allows large-scale prodn. of
CC huDEP-1.
XX
SQ Sequence 5117 BP; 1314 A; 1294 C; 1293 G; 1216 T; 0 other;
```

alignment_scores:

Quality: 508.00 Length: 295
Ratio: 2.579 Gaps: 7
Percent Similarity: 66.780 Percent Identity: 38.305

alignment_block:

US-09-600-358A-4 x AAT06027 ..

Align seg 1/1 to: AAT06027 from: 1 to: 5117

seq name: /net/abss06/SIPDS1/qcadata/hold-geneseq/qeneseqn-emb1/NA2001A.DAT:AAI58418

[illegible]

alignment_scores:		
Quality:	506.50	Length: 592
Ratio:	1.549	Gaps: 21
Percent Similarity:	55.236	Percent Identity: 28.547

alignment_block:

US-09-600-358A-4 x AAI58418

Align seg 1/1 to: AAI58418 from: 1 to: 4798

```

2 AspGlnArgGluIleLeuGlnLys..... 9
1682 CAACAGCAGGAGCTGTGTAAGAGGATGATGAAAAACAACACTGATGAATGT 1731
10 .....PheLeuAspGluAlaGlnSerLysLysIleT 20
1732 GGAGCCCAATCCATCAGATATTTTGGAAACTTATAAGAGGAAGATTG 1781
20 hrLysGluGlu.....PheAlaAsnGluPheLeuLysLeuLysArgGln 34
1782 CTGATGAGGAGAGACTTTTCTCGCTGNAATTCAGAGCATCCCGCGGGTG 1831
35 SerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGluAs 51
1832 TTCAGCAAG.....TTTCCTATAAGGAAGCTCGAAA 1863
51 nAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAspT 68
1864 GCCCTTTAAACCAGAAATAAAACCGTTATGTTGACATTTCTCCTTATGATT 1913
68 yrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSerTyr 84
1914 ATACCGTGTGTGAACCTCTCTGAGATAACGGAGATGAGGCTCAAACTAC 1963
85 IleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIleAl 101
1964 ATAAATGCCACTATATTGATGGTTTCAAGAACCCAGGAATACATTGC 2013
101 aThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetIleT 118
2014 TGCAACAGGTCCTCCAGGATGAAGATGTTGATGATTTCTGGAGGATGATT 2063
118 rpGluTyrSerValLeuIleValMetAlaCysMetGluTyrGluMet 134
2064 GGGACAGAAAGCCACAGTATTATGTCATGTCACCTCGATGTGAAGAAGGA 2113
135 GlyLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGlnLe 151
2114 AACAGGAACAAGTCTGCAGAACTACTGCGCGTCAATGGAAGAGGCCACTCG 2163
151 uGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSerA 168
2164 GGCTTTTGGAGATGTTGTTGTAAGATCAACAGCACACAAAAGATGCCAG 2213
168 sPTyrIleIleArgThrLeu.....LysValLysPheAsnSerGlu 181
2214 ATTACATCTTCAGAAATTCAGATTTGAAACATTTGTAATAAAAAAAGAACAC 2263
182 ThrArgThrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspV 198
2264 TGAAGAGAGAGTGACTCAGATTCAGTTCCACAGCTGGCCAGACCACCGGG 2313
198 alProSerSerIleAspProIleLeuGluLeuIleTrpAspValArgCys 214
2314 TGCCTGAGAGCTCTCACTTGTCTCTCAAACTGAGAAGAGAGATGAATGCC 2363
215 TyrGlnGluAspAspSerValProIleCysIleHisCysSerAlaGlyCy 231
2364 TTCAGCAATTTCTTCAGTGGTCCCATTTGTGGTGCACTGCAGTGTGTTGT 2413
231 sGlyArgThrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuL 248
2414 TGGCGGCACAGAACCTATATCGGAATT.....GATGCCA 2448
248 eutLysAspGlyIleIleProGluAsn...PheSerValPheSerLeuIle 263
2449 TGCTAGAAGGCTTGGAGCCGAGAACAAAGTGGATGTTTATGTTATGTT 2498

```

```

264 ArgGluMetArgThrGlnArgProSerLeuValGlnThrGlnGluGlnTy 280
2499 GTCAGCTTAAGCGCAGACAGATGCTGATGTTCAAGTAGGAGGCCAGTA 2548
280 rGluLeuValTyrAsnAlaValLeuGluLeu.....PheLysArgGlnM 295
2549 CATCTTGATCCATCAGGCTTTGGTGAATACAATCAATCAGTTTGGAGAAACAG 2598
295 etAspValIleArgAspLysHisSerGlyThrGluSerGlnAlaLysHis 311
2599 AAGTGAATTTGCTGTAATTACATCCATATCTACATAACATGAAGAAAGG 2648
312 CysIleProGluLysAsnHisThrLeuGlnAlaAspSerTyrSerProAs 328
2649 GATCCACCCAGTGAGCGCTCTCCACTAGAGCTGAA.....TTCCAGAG 2692
328 nLeuProLys...SerThrThrLysAlaAlaLysMetMetAsnGlnGlnA 344
2693 ACTTCTTCATATAGGAGCTGGAGACACACACATTTGGAATCAAGAAG 2742
344 rgThrLysMetGluIleLysGluSerSerSerPheAspPheArgThrSer 360
2743 AAAATAAAGTAAACAGGAATTTAATGTCATCCCATATGACTATAAAC 2792
361 GluIleSerAlaLysGluLeuValLeu.....HisPr 372
2793 AGAGTGCCACTTAAACATGAGCTGGAATGAGTAAAGAGAGTGAGCATGA 2842
372 oAlaLysSerSerThrSerPheAspPheLeuGluLeuAsnTyrSerPheA 389
2843 TTCAGATGAATCCTCTGTATGATGACAGTGAATTCAGAGGAACCAAGCAAT 2892
389 spLysAsnAlaAspThrThrMetLys...TrpGlnThrLysAlaPhePro 404
2893 ACATCAATGATCTTTTATATGAGTACTGGAACCTGAAAGTATGATGATT 2942
405 IleValGlyGluProLeuGlnLysHisGlnSerLeuAspLeuGlySerLe 421
2943 GCTGCTCAGGACCACCTG...AAGGAGACCATTGCTGACTTTTGGCAGAT 2989
421 uLeuPhe..... 423
2990 GATCTTCCAAAAGAAAAGTCAAAGTTATTGTTATGCTGCAGACAAGTGAAC 3039
424 .....GluGlyCysSer.....AsnSerLysProVal 432
3040 ATGGAGACCAAGGAAATCTGTCTCAGTACTGGGGAGGAAGGAACCAACA 3089
433 AsnAlaAlaGlyArgTyrPheAsnSerLysValProIleThrArgThr.. 448
3090 .....TATGGAGATATTGAAGTTCACCTGACCTGAAAGACACAGA 3124
449 LysSerThrProPheGluLeu.....IleGlnGlnArgGluThrLysG 463
3125 CAAATCTTCAACTTATACCTCTGCTGCTTTTGAAGTCTGAGACATTCACA 3174
463 luValAspSerLysGluAsnPheSerTyr..... 472
3175 GGAAGAGACTCTGAACTGTGTACCACTACCAATATACAACTGGAGTGTG 3224
473 .....LeuGluSerGlnProHisAsp...SerCysPheValGluMetGl 486
3225 GAGCAGCTTCTCGAACCAACCAAGGAATTAATCTATGATTACAGGTGCT 3274
486 nAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyrSerLeup 503
3275 CAAACAAAACCTCCCGAGGAAGATTTCTCTGAAGGGGAC..... 3314
503 roTyrAspSerLysHisGlnIleArgAsnAlaSerAsnValLysHisHis 519
3315 .....AAGCATCAC 3323

```



```

87  aasnPhelLeLysGlyValTyrGlyProLysAlaTyrIleAlaThrGlnG 104
   |||||:||||| ||| |||||:||||| |||||:||||| |||||
2847 CAATATATTGATGGCTACAGAGACCAAGTCAATTCATTGACCAACCAAG 2896
104  LyrProLeuSerThrThrLeuLeuAspPheThrArgMetIleTrpGluTyr 120
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
2897 GTCCCGTTTCATGAACACGATGTATGATTTCTGGAGGATGATTGGCAAGAA 2946
121  SerValLeuIleIleValMetAlaCysMetGluTyrGluMetGlyLysLy 137
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
2947 CAATCTGCTTGCATTGCTGATGCTTACAAATTTAGTTGAGTTGGCCGGT 2996
137  sLysCysGluArgTyrTrpAlaGluProGlyGluMetGlnLeuGluPheG 154
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
2997 TAAATGCTATAAATATTGGCCCTGATACTCAAGTT.....TATG 3037
154  LyrProPheSerValSerCysGluAlaGluLysArgLysSerAspTyrIle 170
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
3038 GTGACTTCAAGTAACTGCTGTAGAAATGGAAACCACTTGCTGAATATGTA 3087
171  IleArg.....ThrLeuLysValLysPheAsnSerGluThrArgThrI1 185
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
3088 GTTAGACATTCACCTCGAAAGGAGGGGTACAAATGAATCCGTGAAGT 3137
185  eTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSerSerI 202
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
3138 TAAACAGTTCCATTTTCACGGGTCGCTGACCATTGAGTGCCCTACCATG 3187
202  leAspProIleLeuGluIleTrpAspValArgCysTyrGlnGluasp 218
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
3188 CPACAGGGCTGCTTTCCTTTATCCGGCGGAGTCAAGTTATCAACCCCTCC 3237
219  AspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgThrG1 235
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
3238 AGTGTGCGCCCATCGTGTGATTCATTCAGTGTGCTGTGGACGACATGG 3287
235  yValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAspGlyI 252
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
3288 CTGTACATTTGTGATGTACATCATGCTAGACATGGCTGAAGAGAGGGTG 3337
252  leIleProGluAsnPheSerValPheSerLeuIleArgGluMetArgThr 268
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
3338 TTGTT.....GATATTTCAATTTGTCAAAAGCCCTTAAGATCT 3375
269  GlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValTyrAs 285
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
3376 CGGCGTATTATATGTCGACAGAGGAAACAGTACATTTTATTTCATGA 3425
285  nAlaValLeuGluLeu..... 290
   |||||:|||||
3426 TGCCATTTTAGAAGCTGCTTATGTGGAGAAACTGCCATACCTGCTGTG 3475
291  ..PheLys...ArgGlnMetaspValIleArg..... 299
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
3476 AATTTAAAGCTGCATATTTTGTATGATGATTAGAATAGACTCCAGACTAAC 3525
300  AspLysHisSerGlyThrGluSerGlnAlaLysHis..... 311
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
3526 TCCTTCATCTCAAGATGAATTCAGACTCTGAATTCAGTCAACCCCTCG 3575
312  ..... 319
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
3576 ACTACAAGCTGAAGACTGCAGTATAGCTGCTGCCA...AGGAACCATG 3622
319  hrLeuGln.....AlaAspSerTyrSerProAsn.....LeuProLys 331
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
3623 ACAAGAACCGTTTCATGGACATGCTGCCACCTGCAGAGATGCTGCTCTTT 3672
332  SerThrThrLysAlaAlaLysMetMetAsnGlnGlnArgThrLysMet... 347
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
3673 TTAATTACAATTGATGGGAGAGCAGTAACTACATCAATGCTGCTCTTAT 3722
348  ....GluIleLysGluSerSerPheAspPheArgThrSerGluIles 363

```

```

   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
3723 GGACAGCTACAGGCAACCAAGCTGCTTTC..... 3750
363  erAlaLysGluGluLeuValLeuHisProAlaLysSerSerThrSerPhe 379
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
3751 .....ATCGTCACACAAATACCCCTCTGCCAAACACACTGTAAAA 3786
380  AspPheLeuGluLeuAsnTyrSerPhe..... 388
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
3787 GACTTCTGGAGATTAGTGTATGATATGGCTGTACCTCCATTGTGATGTT 3836
389  .....AspLysAsnAlaAspThrThrMetLysTrpGlnThrLysA 402
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
3837 AAACGAAGTCGACTTGTCCAGGGCTGCCCTCAGTACTGCCAGAGGAAG 3886
402  laPhe.....ProIleValGlyGluProLeuGlnLysHisGln 414
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
3887 GGATCTACGATATGGCCCATCCAAAGTGAATGATGTCTTTGTTCAATG 3936
415  SerLeuAspLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysPr 431
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
3937 GACTGTGATGTATCAACCGGATTTTAGGATATGCAATCTAACAGACC 3986
431  oValAsnAla.....AlaGlyArgTyrP 439
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
3987 ACAGGAAGTTATCTGATGGTGCAACAGTTTCAGTACCTAGGATGGCTT 4036
439  heAsnSerLysValProIleThrArgThrLysSerThrProPheGluLeu 455
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
4037 CTCATCGAGAGTGCCT.....GGATCCAAAGGTCATCTTGAACACTG 4080
456  ileGlnGlnArgGluThrLysGluValAspSerLysGlu 468
   |||||:||||| |||||:||||| |||||:||||| |||||:|||||
4081 ATACTTCAGGTGAAAGTGGCAGGAGGAATCCAGGAG 4119

```

OM of: US-09-600-358A-4 to: GenEmbl.* out_format : pfs
Date: May 26, 2002 2:12 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=framet.p2n.model -DEV=xlp
-Q/cgn2_1/USFTO.spool/US09600358/runat_24052002_101554_27006/app_query.fasta_1.761
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -GAPOP=6.000
-GAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DEEXT=7.000 -START=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09600358_eCgNI_1_11307
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIME=120 -WARN_TIME=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-600-358A-4
Query length: 692
Database: GenEmbl.*
Database sequences: 1797656
Database length: 187333701
Search time (sec): 3968.550000

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
gb_pat:AX020263	+ 3615.00	5088.29	4.7e-275	2356	AX020263 Sequence 3 from Paten
gb_pr:AF001847	+ 3611.00	5082.64	9.6e-275	2356	AF001847 Homo sapiens lymphoid
gb_pat:AX020261	+ 3573.00	5026.17	1.3e-271	3058	AX020261 Sequence 1 from Paten
gb_pr:AF001846	+ 3569.00	5020.51	2.8e-271	3058	AF001846 Homo sapiens lymphoid
gb_pat:AR149916	+ 3543.50	4982.80	3.5e-269	3580	AR149916 Sequence 1 from Paten
gb_pr:AR150732	+ 2977.50	4185.30	9.2e-225	2877	AR150732 Homo sapiens protein
gb_to:MUSPROTYPH	+ 2356.50	3308.29	6.5e-176	2734	M90388 Mouse protein tyrosine
gb_pr:AL137856	+ 1490.50	2042.33	2.1e-105	143618	AL137856 Human DNA sequence
gb_pr:HUMPTPGL	+ 1157.00	1612.48	1.9e-81	2938	D13380 Human mRNA for protein-
gb_sts:G18337	+ 1157.00	1611.71	2.1e-81	3160	G18337 SWISS1985 Eric D. Green
gb_pr:HUMPTPPEST	+ 1130.50	1576.03	2.0e-79	2676	M93425 Human protein tyrosine
gb_to:MMPTPPEPS	+ 1104.50	1538.13	2.6e-77	2983	X86781 M.musculus mptp-pest ge
gb_pat:AX306312	+ 1058.00	1479.00	5.1e-74	1608	AX306312 Sequence 1063 from Pa
gb_to:MMP19PTP	+ 767.00	1066.48	4.8e-51	1817	D38072 Rat mRNA for protein ty
gb_to:RATRKPTP	+ 748.00	1037.61	2.0e-49	2198	U69673 Rattus norvegicus prote
gb_to:RNU69673	+ 748.00	1037.61	2.0e-49	2226	AR095603 Sequence 1 from Paten
gb_pat:AR154418	+ 736.50	1025.22	9.6e-49	1529	AR154418 Sequence 1 from Paten
gb_to:BC008512	+ 735.50	1023.04	1.3e-48	1643	BC008512 Mus musculus, Similan
gb_to:MMPTPHSC	+ 733.50	1020.98	1.7e-48	1529	U49853 Mus musculus protein ty
gb_to:MMU35124	+ 733.50	1020.78	1.7e-48	1559	U35124 Mus musculus nonrecept
gb_to:MMU52523	+ 727.50	1012.24	5.1e-48	1568	U52523 Mus musculus protein ty
gb_pr:HSBDP1	+ 710.00	981.29	2.7e-46	2810	X79568 H.sapiens BDP1 mRNA for
gb_to:MUSLY5AA	+ 521.00	708.06	4.5e-31	5007	M92933 Mouse lymphocyte common
gb_pr:HSICAR	+ 516.50	703.28	8.2e-31	4315	Y00638 Human mRNA for leukocy
gb_pr:AF173857	+ 515.50	701.20	1.1e-30	4597	Y00662 Human mRNA for T200 leu
gb_to:MUSC57BL6A	+ 513.00	700.47	1.2e-30	4920	AF173857 Xenopus laevis recept
gb_to:DB3203	+ 513.00	710.40	3.3e-31	1391	D49393 Mouse mRNA for protein
gb_to:MUSBTP	+ 513.00	698.60	1.5e-30	4208	D83203 Mus musculus mRNA, one
gb_to:DB3204	+ 513.00	698.07	1.6e-30	4426	D45212 Mouse mRNA for receptio
gb_to:MUSLY5A	+ 513.00	697.67	1.7e-30	4596	D83204 Mus musculus mRNA, one
gb_to:MUSLY5B	+ 512.00	696.12	2.1e-30	4655	M14342 Mouse 200 kD T-cell leu
gb_pat:AB4503	+ 512.00	694.97	2.4e-30	5186	M14343 Mouse Ly-5 (leucocyte-d
gb_pr:HUMPTPES	+ 509.00	698.24	1.6e-30	2562	AB4503 Sequence 118 from Paten
gb_pat:AR109903	+ 508.00	690.24	4.4e-30	4756	D37781 Human mRNA for protein-
gb_pr:AR10886	+ 508.00	689.46	4.9e-30	5117	AR109903 Sequence 1 from Paten
gb_to:RNU40790	+ 508.00	687.22	6.5e-30	6314	U40790 Rattus norvegicus vascu
gb_to:RATLCA1	+ 502.50	685.61	7.9e-30	3541	M10072 Rat leukocyte common ad

gb_ro:RNLICAR

+ 502.50

684.24

9.5e-30

4027

! Y00065 Rat mRNA for leucocy

gb_ov:AB031424

+ 500.50

680.25

1.6e-29

4490

! AB031424 Cyprinus carpio mR

gb_ro:AY026862

+ 498.00

677.08

2.4e-29

4338

! AY026862 Mus musculus brain

gb_ro:AY026861

+ 498.00

677.02

2.4e-29

4365

! AY026861 Mus musculus brain

gb_ro:AF162856

+ 498.00

676.55

2.5e-29

4562

! AF162856 Mus musculus recep

seq_name: gb_pat:AX020263

seq_documentation_block:

LOCUS

AX020263

2356 bp

DNA

linear

PAT 07-SEP-2000

DEFINITION

Sequence 3 from Patent WO9936548.

ACCESSION

AX020263

VERSION

AX020263.1

GI:10044044

KEYWORDS

human.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 2356)

AUTHORS

Roifman,C.M.

TITLE

Human lymphoid protein tyrosine phosphatases

JOURNAL

Patent: WO 9936548-A 3 22-JUL-1999;

FEATURES

HSC RES DEV LP (CA); ROIFMAN CHAIM M (CA)

source

Location/Qualifiers

1..2356

/organism="Homo sapiens"

/db_xref="taxon:9606"

42..2120

/note="unnamed protein product"

/codon_start=1

/protein_id="CAC07769.1"

/db_xref="GI:10044045"

/translation="MDQREILQFLDEAQSKITKEEFANEFLLKRSQTKYKADKTY

PTTVAAENAKNIKRNKRDILPDYRSVELSITSDDESSYINANFINGVVGPKAYIAT

QGPLSTLLDFWRMIWEYSVLIIIVMACMEYEMGKKCKERYWAEPPGEMQLEFGPFVSC

EAEKRSKDIYIRTLKVKFNSETIYQHYKNPDRHDVPSSIDPILLEIWDVRCYQED

DSVPICITCSAGCGRTVICAIVDYTWMLIKDGIIPENSVFSLIREMRTQRPVLVQT

QOEYELVYNAVLFLFKRQMDVIRDKHSGTSEQAQKHCIPEKNHTLQADSYSPNPKSTT

KAAMNQQTQKWEIKESSEDFRTSEISAKELVLHPAKSSTSFDFLELYSFDNA

DTMKWQTKAFPIPLGFLQHQSLDGLSFLFECSSKPNVNAAGRTFNSKVPITRTKS

TPFELIQQRTEKREVDSENFYSLESQPHDSFVEMQAKQVMHVSSAELNYSLPYDSKH

QIRNASNRRHSDSALGVSYIPLVENPYFSSWPPSGTSKMSLDLPEKQDGVFPFS

LLPTSTSLFYSYNSHSSLSNPTNSSLNQESAVLATAPRDEIDPPPLPVRPPE

SFTVVEAGEFSFNPVKSLSSAVKVKIGTSLEWGGTSEPKKFDSDVILRPSKSVKLRS

PKSGKNFSL"

BASE COUNT

750 a

466 c

466 g

674 t

ORIGIN

alignment_scores:

Quality: 3615.00

Ratio: 5.224

Percent Similarity: 100.000

Percent Identity: 100.000

alignment_block:

US-09-600-358A-4 x AX020263

Align seg 1/1 to: AX020263 from: 1 to: 2356

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17

|||||

42 ATGGACCAAGAGAAATTCGAGAAATTCCTGGATGAGGCCAAGCAA 91

17 slsyleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArgG 34

|||||

92 GAAATTTACTAAGAGAGATTTCGCAATGAATTTCTGAAGCTGAAAGGC 141

34 lnSerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50

|||||

142 AATCTACCAAGTACAGGACAGCAAAACCTATCCTACAACTGTGCTGAG 191

51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67

|||||

192 AATGCCAAGAATAATCAAGAAAAACAGATATAAGGATATTTTGGCCCTATGA 241
67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSer 84
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
242 TTATAGCCGGGTAGAACTATCCCTGATGAACCTCTGATGAGGATTCAGCT 291
84 yTleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
292 ACATCAATGCCAACTTCATTAAAGGAGTTTATGACCCCAAGGCTATATT 341
101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetII 117
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
342 GCCACCCAGGGCTCTTATCTACAACCCCTCTGSACTTCGGAGGATGAT 391
117 eTTPGluTyrSerValLeuIleIleValMetAlaCysMetGluTyrGlu 134
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
392 TTGGGAATATAGTGTCTCTTATCATCTTATGCGCATGCGATGGATGAAA 441
134 etGlyLysLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
442 TGGGAAAAAGAAAGTGTGAGCGCTACTGGGCTGAGCCGAGGAGATGCG 491
151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
492 CTGGAATTTGGCCCTTTCTCTGTATCTCTGTAAAGCTGAAAAAAGGAATC 541
167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArg 184
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
542 TGATATATATATCAGGACTCTAAAGTTTAAGTTCAATAGTGAACCTCGAA 591
184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSer 200
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
592 CTATCTACCAGTTTCATTACAAGAAATGGCCAGACCATGATGATACCTTCA 641
201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnG 217
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
642 TCTATAGACCCCTATTCTGAGCTCATCTGGGATGTACGTTGTACCAGA 691
217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArg 234
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
692 GGATGACAGTGTCCCATATGCAATCTACTGCAAGTGTGGCTGTGGAAGA 741
234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
742 CTGGTGTATTGTGGCTATTGTTCATTATACATGGATGTTGCTAAAGAT 791
251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
792 GGGATAATTCTGAGAACTTCAGTGTTCAGTTCAGTTTGATCCGGGAAATCG 841
267 qThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValT 284
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
842 GACACAGGGCTTCATTAGTTCAACGCGAGGAACAATATGAACCTGCTCT 891
284 yRAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
892 ACAATGCTGTATTAGAATAATTAAGACAGATGGATGTTATCAGAGAT 941
301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
942 AAACATTTCTGGACAGAGAGTCAAGCAAGCAATGTTATTCCTTGAGAAAA 991
317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
992 TCACACTCTCCACGACAGCTCTTATCTCTTAATTTACCAAAAGTACCA 1041
334 hrLysAlaAlaLysMetMetAsnGlnGlnArgThrLysMetGluIleLys 350
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
1042 CAAAAGCAGCAAAATGATGAACCAACAAAGGACAAAAATGGAATCAAA 1091
351 GluSerSerPheAspPheArgThrSerGluIleSerAlaLysGluG 367
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
1092 GAATCTTCTTCTTGTAGCTTTAGGACTTCTGAAATAAGTGCAAAAAGAGA 1141

367 uLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGluL 384
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
1142 GCTAGTTTGGCACCTTGCTAAATCAAGCACTTCTTTTGACTTCTTGAGC 1191
384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnThr 400
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
1192 TAAATTACAGTTTTTGACAAAATGCTGACAAACCATGAATGGCAGACA 1241
401 LysAlaPheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAs 417
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
1242 AAGCATTTCCAAATAGTTGGGAGCCCTTTCAGAAGCATCAAAAGTTTGA 1291
417 pLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsnA 434
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
1292 TTTGGGCTCTCTTTTGTGAGGGATGTTCTAATCTTAAACCTGTAATG 1341
434 laAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSer 450
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
1342 CAGCAGGAAGATATTTTAATTCAAAGGTGCCAATAACACGAGCAATCA 1391
451 ThrProPheGluLeuIleGlnArgGluThrLysGluValAspSerLy 467
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
1392 ACTCCTTTTGAATGTACAGCAGAGAGAAACCAAGGAGGTGGACAGCAA 1441
467 sGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheValG 484
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
1442 GGAACACTTTTCTATTGGAATCTCAACCCACATGATCTTGTGTTGTAG 1491
484 luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
1492 AGATGCGAGGCTCAAAAAGTAAATGCATGTTTCTTCAGCAGAACTGAATAT 1541
501 SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnValLy 517
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
1542 TCCTGTCATATGACTTAAACACCAAAATACGTAATGCCTCTAATGTAAA 1591
517 sHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValG 534
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
1592 GCACCATGACTCTAGTCTCTTGGTGTATATCTTACATACCTTTAGTGG 1641
534 luAsnProTyrPheSerSerTrpProSerGlyThrSerSerLysMet 550
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
1642 AAAATCCTTATTTTTCATCATGCGCTCCAAAGTGTACCACTTCTAAGATG 1691
551 SerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerLe 567
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
1692 TCTCTTGATTTACTGAGAAGCAAGATGGAACGTGTTTTTCTCTCTCTCT 1741
567 uLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSerS 584
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
1742 GTTGCCAAACATCCTCTACATCCCTTCTCTTATTACAATTCACATAGTT 1791
584 erLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnGlu 600
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
1792 CTTTATCAGTGAATTCCTCAACCAATATTTCTCCTACATTTTGAACCCAGAG 1841
601 SerAlaValLeuAlaThrAlaProArgIleAspAspGluIleProProPr 617
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
1842 TCAGCTGTACTAGCACTGCTCCAAGATAGATGAATATCCCTCCCTCC 1891
617 oLeuProValArgThrProGluSerPheIleValValGluGluAlaGlyG 634
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
1892 ACTTCTGTACGGACACCTGAATCATTTATTGTGTTGAGGAAGCTGGAG 1941
634 luPheSerProAsnValProLysSerLeuSerSerAlaValLysValLys 650
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
1942 AATTCTCACCAAAATGTTCCCAAAATCCTTATCTCAGCTGTGAAGGTAAA 1991
651 IleGlyThrSerLeuGluTyrGlyGlyThrSerGluProLysLysPheAs 667
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
1992 ATTGGAACATCAGTGGAAATGGGGTGGGAACATCTGACCAAGAAATTTGA 2041

667 pAspServAlileuArgProSerLysSerValLysLeuArgSerProL 684
 |||||
 2042 TGACTCTGTGATCTTAGACCAAGCAGAGTGTAAACTCCGAGTCCCTA 2091
 684 ySserGlyLysAsnPheSerTrpLeu 692
 2092 AATCAGGTAAATTTCTCTGTGGCTT 2117

seq_name: gb_pr:AF001847

seq_documentation_block:
 LOCUS AF001847 2356 bp mRNA linear PRI 01-OCT-2001
 DEFINITION Homo sapiens lymphoid phosphatase Lyp2 mRNA, complete cds.

ACCESSION AF001847

VERSION AF001847.1 GI:4100633

KEYWORDS

SOURCE

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

1 (bases 1 to 2356)
 Cohen, S., Dadi, H., Shaoul, E., Sharfe, N. and Roifman, C.M.
 Cloning and characterization of a lymphoid-specific, inducible
 human protein tyrosine phosphatase, Lyp

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (30-APR-1997) Immunology and Allergy, The Hospital For
 Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8,
 Canada

FEATURES

source

1. .2356

/organism="Homo sapiens"

/db_xref="taxon:9606"

42. .2120

/note="splice variant of lymphoid phosphatase Lyp1"

/codon_start=1

/product="lymphoid phosphatase Lyp2"

/protein_id="AAD00905.1"

/db_xref="GI:4100634"

/translation="MDQREILQKFLDEAQSKEITKEEFANFEFLKLRQSTKYKADKY

PTTVAENAKIKKRYKIDILPYDSRVLSLITSDSSINAFIKGVGPRAYIAT

QGPLSTILLDFRWMEYSVLIVMACMEYEMGRKKCEKRWAEPEGMQLFEPFVSVC

EAERKSDYIIRTLKVFNSERTIYQFYHKNPMDHPDPSIDPILDLIWDVRCYQED

DSVPICIHCSAGCGRTGVCIAVDYTWMLKDGIPENFVSFLIREMRTQRLSVDT

QEOVELYNVAVLELFKRMQDVIRDKHSGTSEQAOKHCIPKRNHTLQADSYSNLPKSTT

KAAMMMNQRTKMEIKSSSDFDPTSEISAKELVLPKASSTSFDFLELYNYSFDKNA

DTYMKWQTKAPPVIGPELQKHQSLDLSLLEFGCSNKPYNAGRYFNKVPITRTKS

TPFLIQORETKEDVDSKENSYSLESQPHDSCFVEMQAQKVMHVSSAELNYSLPYDSKH

QIRNASNVKHHSDSALGVYSIPLVENPYFSSWPPSGTSSKMSLDLPKQDGVFPSS

LLPTSSLSFSYSNSHDSLSNPTNISLNLQESAVLATAPRIDDEIPPLPVRTPE

SFIVVEEAGFSPNPKVSLSSAVKVIKGTSLWEGTSEPKFDDSVILRPSSKVKLRS

PKSGKNFSWL"

BASE COUNT 750 a 466 c 466 g 674 t

ORIGIN

alignment_scores:

Quality: 3611.00 Length: 692

Ratio: 5.218 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 99.855

alignment_block:

US-09-600-358A-4 x AF001847

Align seg 1/1 to: AF001847 from: 1 to: 2356

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17

|||||

42 ATGGACCAAGAGAAATTCGACAGAGTTCTCTGGATGAGGCCCAAGCAA 91

17 slystleThrLysGluLupheAlaAsnGluPheLeuLysLeuLysArgG 34
 |||||
 92 GAAATTTACTAAGAGAGAGTTTGGCAATGAATTTCTGAAGCTGAAAGGC 141
 34 InSerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50
 |||||
 142 AATCTACCAAGTACAAAGGCAGACAAAACCTATCTACAACTGTGGCTGAG 191
 51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
 |||||
 192 AATGCCAAGAATATCAAGAAAAACACATATAAGGATATTTGCCCTATGA 241
 67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSerT 84
 |||||
 242 TTATAGCCGGGTAGAACTATCCCTGATAACCTCTGTATGAGGATTCAGCT 291
 84 yTleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
 |||||
 292 ACATCAATGCCAACTTCATTAAGGGAGTTTATGGACCCCAAGGCTTATAT 341
 101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetil 117
 |||||
 342 GCCACCCAGGGCTCTTATCTACAACCCCTCTGGACTTCTGGAGGATGAT 391
 117 eTrpGluTyrSerValLeuIleValMetAlaCysMetGluTyrGluM 134
 |||||
 392 TTGGGAATATAGTGTCTTATCATTTATGGCATGCATGGAGATGAGAA 441
 134 eTgLyLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
 |||||
 442 TGGGAAAGAAAAGTGTGAGCGCTACTGGGCTGAGCCAGAGAGATGCAG 491
 151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
 |||||
 492 CTGGAATTTGGCCCTTCTCTGTATCTGTGAAGCTGAAAAAAGGAAATC 541
 167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
 |||||
 542 TGATATATAAATCAGGACTCTAAAAGTTAAGTTCAATAGTGAACCTCGAA 591
 184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProser 200
 |||||
 592 CTATCTACCAAGTTTCTATTAAGAATTTGCCAGACCATCATGTACCTTCA 641
 201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnG 217
 |||||
 642 TCTATAGACCTTATCTTGGCTCATCTGGGATGTACGTTGTACCAAGA 691
 217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
 |||||
 692 GGATGACAGTGTTCATCATGATGATGATGATGATGATGATGATGATGAT 741
 234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
 |||||
 742 CTGGTGTATTTTGTCTATTGTTGATTTATACATGGATGTTGCTAAAAGAT 791
 251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetar 267
 |||||
 792 GGGATAAATTCCTGAGAACTTCAGTGTGTTTTCAGTTTTCATCCGGGAAATGCG 841
 267 gThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValT 284
 |||||
 842 GACACAGAGGCGCTTATAGTATCAACACGAGGAGCAACAATATGAATGGTCT 891
 284 yTAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
 |||||
 892 ACAATGCTGTATTAGAACTATTATTAAGACACATGGATGTTATCAGAGAT 941
 301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
 |||||
 942 AAACATCTCTGGAACAGAGAGTCAAGCAAGCATTTGTTCTCTGAGAAAAA 991

```
317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThr 334
|||||
992 TCACACTCTCCACGACACTCTATTCTCTCTAATTTACCAAAAAGTACCA 1041
|||||
334 hrLysAlaAlaLysMetMetAsnGlnGlnArgThrLysMetGluLeLys 350
|||||
1042 CAAAAGCAGCAAAAATGATGAACCAACAAGGAGCAAAAATGGAATCAAA 1091
|||||
351 GluSerSerPheAspPheArgThrSerGluLeuSerAlaLysGluG1 367
|||||
1092 GAATCTTCTCTCTTGAATTTAGGACTTCTGAATAAGTGCAAAAAGAA 1141
|||||
367 uLeuValLeuHisProAlaLysSerThrSerPheAspPheLeuGluL 384
|||||
1142 GCTAGTTTTCACCTGCTAAATCAAGCACTTCTTTTGACTTCTGGAGC 1191
|||||
384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTyrGlnThr 400
|||||
1192 TAAATTACAGTTTTGACAAAAATGCTGACACCAACCATGAATGGCAGACA 1241
|||||
401 LysAlaPheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAs 417
|||||
1242 AAGGCATTTCCAAATAGTTGGGAGCCCTTCAGAAAGCATCAAAAGTTGGA 1291
|||||
417 pLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsnA 434
|||||
1292 TTTGGGCTCTCTTTGTTGAGGGATGTTCTAATCTTAACCTGTAATG 1341
|||||
434 laAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSer 450
|||||
1342 CAGCAGGAAGATATTTAATTCAAAGGTGGCAATTAACACGAGCAACAAATCA 1391
|||||
451 ThrProPheGluLeuIleGlnArgGluThrLysGluValAspSerLy 467
|||||
1392 ACTCTTTTGAATTTGATACAGCAGAGAGAACCAAGGAGGTGGACAA 1441
|||||
467 sGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheValG 484
|||||
1442 GGAACACTTTCTTATTTGGAATCTCAACACATGATTCGTGTTGTAG 1491
|||||
484 luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
|||||
1492 AGATGCAGGCTCAAAAAGTATGATGATGTTCTTCAGCAGAACTGAATAT 1541
|||||
501 SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnValLy 517
|||||
1542 TCACTGCCATGATGACTCTAAACACCAATAGTAACTGCTCTAATGTAAA 1591
|||||
517 sHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValG 534
|||||
1592 GCACCATGACTCTAGTCTCTTGGTGATATTTCTTACATACCTTTAGTGG 1641
|||||
534 luAsnProTyrPheSerSerTrpProProSerGlyThrSerSerLysMet 550
|||||
1642 AAAATCCTTATTTTTCATCATGGCTCCCAAGTGTACCACTCTTAAGATG 1691
|||||
551 SerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerLe 567
|||||
1692 TCTCTGATTTTACCTGAGACAGATGGAAGTCTTTTCTCTCTCTCTCT 1741
|||||
567 uLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSerS 584
|||||
1742 GTTGCCACATCTCTACATCCCTCTCTCTTATTACAAATTCACATGATT 1791
|||||
584 erLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnGlu 600
|||||
1792 CTTTATCACTGAATTTCTCAACCAATATTTCTCTCACTATTGAACAGGAG 1841
|||||
601 SerAlaValLeuAlaThrAlaProArgIleAspAspGluIleProProPr 617
|||||
1842 TCAGTGTACTAGCAACTGCTCCAGGATAGATGATGAATCCCCCTCC 1891
|||||
617 oLeuProValArgThrProGluSerPheIleValIleGluAlaGlyG 634
```

```
|||||
1892 ACTTCTCTGTACGACACCTGAATCATTTATTGTGTTGAGGAAGCTGGAG 1941
|||||
634 luPheSerProAsnValProLysSerLeuSerSerAlaValLysValLys 650
|||||
1942 AATTCTCACCAAAATGTCCCAAAATCTCTATCTCAGCTGTGAAGGTAATA 1991
|||||
651 IleGlyThrSerLeuGluTyrGlyGlyThrSerGluProLysLysPheAs 667
|||||
1992 ATTGGAAACATCACTGGATGGGTGGACATCTGAACCAAGAATTTGA 2041
|||||
667 pAspSerValIleLeuArgProSerLysSerValLysLeuArgSerProL 684
|||||
2042 TGACTCTGTGATCTTAGACCAAGCAAGAGTGTAAAACCTCCGAAGTCCTA 2091
|||||
684 ySSerGlyLysAsnPheSerTrpLeu 692
|||||
2092 AATCAGGTAAAAATTTCTCTTGGCTT 2117
|||||
seq_name: gb_pat:AX020261
seq_documentation_block:
LOCUS AX020261 3058 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 1 from Patent WO9936548.
ACCESSION AX020261
VERSION AX020261.1 GI:10044042
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 3058)
AUTHORS Roifman, C. M.
TITLE Human lymphoid protein tyrosine phosphatases
JOURNAL Patent: WO 9936548-A 1 22-JUL-1999;
HSC RES DEV LP (CA); ROIFMAN CHAIM M (CA)
FEATURES
source
1..3058
/organism="Homo sapiens"
/db_xref="taxon:9606"
42..2468
/note="unnamed protein product"
/codon_start=1
/db_xref="GI:10044043"
/translation="MDQELQKFLDEAQSKKITKEEFANFLKLRSTKYKADKTY
PTVAENAKNKKRYKDLIDYDSRVLSLITSDSSYINANFIKGVYGPKAVIAT
EAKRKSQDIIRTLKVFNSERTIYQHYKNPDHPDPSIDPLELIDWDVQSOED
DSVPICIHCSAGCGRTGIVCAIVDYTWMLLDGIIIPENFSVFLIREMRTQPSLVQT
QFOVELVYNAVLELFPKROWDVRDKHSGTESQAKHCIPKXNHTLQADSYSPNLPKSTT
KAAKMNQORTKMEIKESSDFRTSEISAKEELVLPKAKSSTSFDFLELNYSPDKNA
DTTMKWTQAFPIVGEPLQKQSLDGLSGLLFEGCSNKPVRNAGRFNPKVPIITKKS
TFELIQORETKVEDSKENFSYLESQPHDSCFVEMQAOKVMHVSAELNYSILPDSKH
QIRNASNVKHHDSALGVSYIPLVNFYFSWPPSGTSSKMSLDLPKQDGTVPFS
LIPTSTSLFSYNSHSLNSPTNISSLNQESAVLATAPRIDDEIPPLPVPRTPE
SFIVVEAGEFSPNPVKLSAVKVIKIGTSLSEWGTSTPKFPDDSVILRPKSKVLRS
PKSELHODRSPPPPLPERTLESFLLADECMQAOSIETYSYPTDWENSTSSKQTL
KTPGKSTFRSLKILRNKKKSICNCSPPNKPAPESVQSNSSSFLNFGFANRFSKPKG
PRNPPTTWN1"
BASE COUNT 1015 a 592 c 553 g 898 t
ORIGIN
alignment_scores:
Quality: 3573.00 Length: 685
Ratio: 5.216 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-600-358A-4 x AX020261
Align seg 1/1 to: AX020261 from: 1 to: 3058
```


1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerly 17
|||||
42 ATGGACCAAGAGAAATTCGAGAGATTCCTGGATGAGGCCCAAGACAA 91
|||||
17 slyIleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArg 34
|||||
92 GAAATTTACTAAGAGAGAGTTTGGCAATGAATTTCTGAAGCTGAAAAGC 141
|||||
34 InSerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50
|||||
142 AATCTACCAAGTACAAAGGAGACAAACCTATCCTACAACTGTGGCTGAG 191
|||||
51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
|||||
192 ANTGCCAAGAATATCAAGAAACAGATATAAGGATATTTGCCCTATGA 241
|||||
67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSer 84
|||||
242 TTATAGCCGGGTAGAACTATCCCTGATACCTCTGATGAGGATTCAGCT 291
|||||
84 yIleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
|||||
292 ACATCAATGCCAACCTTCATTAAGGAGTTTATGGACCCCAAGGCTTATAT 341
|||||
101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTyrArgMetI 117
|||||
342 GCCACCCAGGGTCCCTTATCTACACCCCTCTGGACTTCGAGGATGAT 391
|||||
117 eTrrGluTyrSerValLeuIleValMetAlaCysMetGluTyrGluM 134
|||||
392 TTGGGAATATAGTCTCTTATCATTTGTTATGGCATGCATGGAGTATGAA 441
|||||
134 eGlyLysLysLysCysGluArgTyrTrrAlaGluProGlyGluMetGln 150
|||||
442 TGGGAAGAAAGAAAGTGTAGCGCTACTGGGCTGAGCCAGAGAGATGCAG 491
|||||
151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
|||||
492 CTGGAAATTTGGCCCTTCTCTGTATCTCTGTGAAGCTGAAAAGGAATC 541
|||||
167 rAspTyrIleLeuArgThrLeuLysValLysPheAsnSerGluThrArg 184
|||||
542 TGAATATATATATACAGACTCTAAAAGTTAAGTTCAATAGTGAACCTCGAA 591
|||||
184 hrIleTyrGlnPheHisTyrLysAsnTrrProAspHisAspValProser 200
|||||
592 CTATCTACCAAGTTTCATTACAAGAATTGGCCAGACCATGATGACCTTCA 641
|||||
201 SerIleAspProIleLeuGluLeuIleTrrAspValArgCysTyrGlnGl 217
|||||
642 TCTATAGACCCCTATTCCTGGAGCTCATCTGGGATGACGTTGTACCAAGA 691
|||||
217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
|||||
692 GGATGACAGTGTTCACATATGCAATTCACCTGCAGTGTGCTGTGAAGA 741
|||||
234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
|||||
742 CTGGTGTATTTGTGCTATTTGTGATATATACATGGATGTTGCTAAAAGAT 791
|||||
251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267
|||||
792 GGGTAATTCCTGAGAATTCAGTGTTCAGTTTGTATCCGGGAAATGCG 841
|||||
267 gThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValT 284
|||||
842 GACACAGAGGCCCTTCATTAGTTCAACACGACGAGCAACAATATGAACCTGGTCT 891
|||||
284 yrAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
|||||
892 ACAATGCTGTATTAGAACTATTTAAGAGACAGATGGATGTTATCAGAGAT 941
|||||

301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
|||||
942 AAACATTCCTGGACACAGAGACTCAAGCAAGCAATGTATTCTCTGAGAAAA 991
|||||
317 nhIsthrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
|||||
992 TCACACTCTCCAAGCAGACTTATTCTCTAATTTACCAAAAAGTACCA 1041
|||||
334 hrLysAlaAlaLysMetMetAsnGlnArgThrLysMetGluIleLys 350
|||||
1042 CAAAAGCAGCAAAAATGATGAACCAACAAAGGACAAAAATGGAATCAAA 1091
|||||
351 GluSerSerSerPheAspPheArgThrSerGluIleSerAlaLysGluG 367
|||||
1092 GAATCTCTCTCTTTCACCTTAGGACTTCTGAAATAAGTGCACAAAGAGA 1141
|||||
367 uIeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGlu 384
|||||
1142 GCTAGTCTTTCACCTGCTAAATCAAGCACTCTCTTTGACTTTCTGGAGC 1191
|||||
384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrrpGlnThr 400
|||||
1192 TAAATTTACAGTTTTCACAAAAATGCTGACACCAACCAATGAAATGGCAGACA 1241
|||||
401 LysAlaPheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAs 417
|||||
1242 AAGGCATTTCCAATAGTTGGGAGGCTCTTCAGAGCATCAAGTTTGA 1291
|||||
417 pLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsnA 434
|||||
1292 TTTGGGCTCTCTTTGTTGAGGATGTTCTAATCTAAACCTGTAATG 1341
|||||
434 laAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSer 450
|||||
1342 CAGCAGGAAGATATTTAATCAAGGTGCCAATAACACGACCAAAATCA 1391
|||||
451 ThrProPheGluLeuIleGlnArgGluThrLysGluValAspSerly 467
|||||
1392 ACTCTCTTTTGAATTCATCAGCAGAGAGAAACCAAGGAGGTGGACAGCAA 1441
|||||
467 sGluAsnPheSerTyrLeuGluSerGlnProHisaspSerCysPheValG 484
|||||
1442 GCAAAACCTTTCTTATTTGGAATCTCAACACATGATCTCTGTTGTAG 1491
|||||
484 luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
|||||
1492 AGATCAGGCTCAAAAAGTAATGCATGTTTCTTCAGCAGAACTGAATAT 1541
|||||
501 SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnVally 517
|||||
1542 TCACGTCCCATATGACTCTAAACACCAATACGTAATGCCTCTAATGTAA 1591
|||||
517 shHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValG 534
|||||
1592 GCACCATGACTCTAGTGTCTTGGTGATATATCTTACATACCTTTAGTGG 1641
|||||
534 luAsnProTyrPheSerSerTrrProProSerGlyThrSerSerLysMet 550
|||||
1642 AAAATCCTTATTTTCATCATGGCTCCCAAGTGGTACCAGTTCTAAGATG 1691
|||||
551 SerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerle 567
|||||
1692 TCTCTTGATTTACCTGAGAGCAAGATGGAAGTGTTCCTCTCTCTCT 1741
|||||
567 uLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSerS 584
|||||
1742 GTTGCCAAACATCCTCTACATCCCTCTCTCTTATTACAAATTCACATAGTT 1791
|||||
584 erLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnGlu 600
|||||
1792 CTTTATCACTGAATTCACCAACCAATATTTCTCCTACTATTGAACGAGGAG 1841
|||||
601 SerAlaValLeuAlaThrAlaProArgIleAspAspGluIleProPr 617
|||||

|||||
1842 TCAGCTGTACTAGCAACTGTCTCCAGGATAGATGATAAATCCCCCTCC 1891
617 oLeuProValArgThrProGluSerPheIleValValGluGluAlaGlyG 634
1892 ACTTCCCTGTACGGACACTGAATCATTTATGTGTGAGGAAGCTGGAG 1941
634 luPheSerProAsnValProLysSerLeuSerSerAlaValLysVallys 650
1942 AATTCTCACCAAAATGTTCCCAATCCTTATCTCAGCTGTGAAGGTAAAA 1991
651 lIleGlyThrSerLeuGluTrpGlyGlyThrSerGluProLysLysPheAs 667
1992 ATTGGAACATCAGTCGAATGGGTGGAACATCTCGAACCAAGAAATTTGA 2041
667 pAspSerValIleLeuArgProSerLysSerValLysLeuArgSerProL 684
2042 TGACTCTGTGATACTTGAACCAAGAGGTGAATAACTCGGAAGTCCCTA 2091
684 ySSer 685
|||||
2092 AATCA 2096

seq_name: gb_pr:AF001846

seq_documentation_block:

LOCUS AF001846 3058 bp mRNA linear PRI 01-OCT-2001
DEFINITION Homo sapiens lymphoid phosphatase LyP1 mRNA, complete cds.

ACCESSION AF001846

VERSION AF001846.1 GI:4100631

KEYWORDS

SOURCE

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 3058)

Cohen,S., Dadi,H., Shaoul,E., Sharfe,N. and Roifman,C.M.

Cloning and characterization of a lymphoid-specific, inducible

human protein tyrosine phosphatase, Lyp

Blood 93 (6), 2013-2024 (1999)

99168989

10086674

2 (bases 1 to 3058)

Roifman,C.M.

Direct Submission

Submitted (30-APR-1997) Immunology and Allergy, The Hospital For

Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8,

Canada

Location/Qualifiers

1..3058

/organism="Homo sapiens"

/db_xref="taxon:9606"

/cell_type="thymocytes"

42..2468

/codon_start=1

/product="lymphoid phosphatase LyP1"

/protein_id="AAD0904.1"

/db_xref="GI:4100632"

/translation="MDQREILOKFLDEAOSKKITKEEFANFLKLRQSTKYKADKTY

PTVAENAKNKKRYKIDLPDYSRVLSLITSDSDSYINANFIKVGYPKAYIAT

QGPLSTLLDFRMWIEYSLVLIIVMACMEYEMGKKKERYWAEPEGMELEFGPFYSVC

EAEKKSDYIIRTLKVFNSERTIYQFHYKNWPDHVPSSIDPILFLIWDVRCVQED

DSVPICIHCSACGRGTGVCIAVDYTWMLLKDGIIPENSVFSLRDEMRTPRSLVOT

LPSTSLVNAVLELFRKQNDVIRDKHSGTFSQAKHCIPENKHTLOADSYSPNLPKSTT

QAKAMNQRTKWEIKESSDFPRTSEISAKBELVLHPAKSTSFDFLELNTSFDKNA

DTMKWQTKAFPIVGLPEQKHSLDGLSLLFEGCSNKPVNAAGRYFNSKVPITRTKS

TFPELQQRTEKVSKEFNSYFLEQPHDSCFVMOAQKVMHVSAAELNYSLPYDSKH

QIRNASVHRHDSALGVSYIPLVENPYSFSSWPGSTSKMSLDLPEKQDGTVPFS

LIPTSSVYNNHSDLSLNSPTNISLLNOSAVLATAPRIDDEIPPLPVPTPE

STIVVEAGEFSPNPKSLSSAVKVKIGTSLWGGTSEPKFDDSVILRPSKVLRS

PKSELHQDRSSPPPLPERTLESFFLADECHQAOISIEYTSYPTDMENSSSKOTL

KTPGSKFSKSLKILLRNMKKSICNSCPPNPKPAESVQSNSSSFLNFGFANRFSRPGK

PRNPPTWNI"

FEATURES

source

CDS

BASE COUNT 1015 a 592 c 553 g 898 t
ORIGIN

alignment_scores:
Quality: 3569.00 Length: 685
Ratio: 5.210 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.854

alignment_block:
US-09-600-358A-4 x AF001846 ..

Align seg 1/1 to: AF001846 from: 1 to: 3058

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17
|||||
42 ATGAGCCAAAGAGAAATTCGCAGAAAGTTCTGTGATGAGGCCCAAGCAA 91
17 sLysIleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArgG 34
|||||
92 GAAAATTACTAAAGAGAGGTTTGCCAAATGAATTTCTGAAGCTGAAAGGC 141
34 lnSerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50
|||||
142 AATCTACCAAGTACAGGCGAGACAAAACCTATCTACAACCTGTGGCTGAG 191
51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
|||||
192 ATGCCAAGATATCAAGAAAAACAGATATAGGATATTTGCCCTATGA 241
67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSert 84
|||||
242 TTATAGCGGGTAGAAGCTATCCCTGATAACCTCTGATGAGGATCCAGCT 291
84 yrIleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
|||||
292 ACATCAATGCCAACTTCAATTAAGGAGTTTATGACCCCAAGGCTTATAT 341
101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTyrArgMetII 117
|||||
342 GCCACCCAGGTCCTTTATCTACAAACCTCCTGGACTTCTGGAGGATGAT 391
117 eTrpGluTyrSerValLeuIleIleValMetAlaCysMetGluTyrGluM 134
|||||
392 TTGGGAATATAGTGTCTTATCATTTGTTATGGCATGCGATGAGTATGAAA 441
134 etGlyLysLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
|||||
442 TGGGAAAAAAGTGTGAGCGCTACTGGGCTGAGCCAGGAGATGCGAG 491
151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
|||||
492 CTGGAATTTGGCCCTTCTCTGTATCTCTGTAAGCTGAAAAAGGAAATC 541
167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
|||||
542 TGATTATATAATCAGGACTCTAAAAAGTTAAGTTCAAATAGTGAACCTCGAA 591
184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSer 200
|||||
592 CTATCTCCACAGTTTCATTACAGAAATTGGCCAGACCATGATGATACCTTCA 641
201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnGl 217
|||||
642 TCTATAGACCCCTATTTCTTGACCTCATCTGGGATGTACGTTGTTACCAAGA 691
217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
|||||
692 GGATGACAGTGTTCACATATGCATTCACCTGCTGCTGCTGCTGCTGCTGCTG 741
234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysasp 250
|||||
742 CTGGTGTATTATTGTGCTATTGTGATTATACATGATGATGTTGCTAAAGAT 791

```
251 GlyIleIleProGluAsnPhSerValPheSerLeuIleArgGluMetAr 267
|||||
792 GGGATAATTCCTGAGAACTTCAGTGTTCAGTTTGATCGGGAATCGG 841
|||||
267 gThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValT 284
|||||
842 GACACAGAGGCTTCATTAGTTCAAAGCAGGAAACAATATGAACGTGGTCT 891
|||||
284 yAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
|||||
892 ACAATGCTGTATTAGAATAATTAAGAGACAGATGGATGTTATCAGAGAT 941
|||||
301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
|||||
942 AAACATTCGACAGAGAGCTCAAGCAAGCAATGTTATCCTCGAGAAAA 991
|||||
317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
|||||
992 TCACACTCTCCAGCAGACTCTTATCTCCTAAATTTACCAAAAAGTACCA 1041
|||||
334 hrLysAlaAlaLysMetMetAsnGlnGlnArgThrLysMetGluIleLys 350
|||||
1042 CAAAGCAGCAAAAATGATGAACCAACAAAGGACAAATAAGTGAATCAAA 1091
|||||
351 GluSerSerPheAspPheArgThrSerGluIleSerAlaLysGluG1 367
|||||
1092 GAATCTCTTCCTTTGACTTTAGGACTCTCGAATAAGTGCNAAAAGAGA 1141
|||||
367 uLeuValLeuHisProLysSerSerThrSerPheAspPheLeuGluL 384
|||||
1142 GCTAGTTTTCACCCCTGCTAAATCAAGCAGCTCTCTTTGACTTCTCGGAGC 1191
|||||
384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnThr 400
|||||
1192 TAAATTACAGTTTTCACAAAATGCTGACACCAACCAATGAATGGCAGACA 1241
|||||
401 LysAlaPheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAs 417
|||||
1242 AAGCATTTCCATAGTTGGGAGCCCTTCACAGAACATCAAGTTTGGGA 1291
|||||
417 pLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsnA 434
|||||
1292 TTTGGGCTCTCTTTGTTGTTGAGGAGTCTCTAAATCTAAACCTGTAATG 1341
|||||
434 laAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSer 450
|||||
1342 CAGCAGGAAGATATTTTAATTCAAAGGTGCCAATAACACGACCAAAATCA 1391
|||||
451 ThrProPheGluLeuIleGlnArgGluThrLysGluValAspSerLy 467
|||||
1392 ACTCCTTTTGAATGTATACAGCAGAGAGAAACCAAGGAGGTGGACACAA 1441
|||||
467 sGluAsnPhSerTyrLeuGluSerGlnProHisAspSerCysPheValG 484
|||||
1442 GGAACACTTTTCTTATTTGGAATCTCAACCAACATGATCTTGTGTTGTAG 1491
|||||
484 luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
|||||
1492 AGATGCAGGCTCAAAAAGTAAATGCATGCTTCTTCAGCAGAACTGAATAT 1541
|||||
501 SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnValLy 517
|||||
1542 TCACTGCCATATGACTCTAACACCAAAATAGTAATGCCCTCTAATGTAAA 1591
|||||
517 sHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValG 534
|||||
1592 GCACCATGACTCTAGTCTCTTGGTGTATATCTTACATACCTTTAGTGG 1641
|||||
534 luAsnProTyrPheSerSerTrpProSerGlyThrSerLysMet 550
|||||
1642 AAAATCCTTATTTTCATCATGCGCTCCAAAGTGGTACCACTTCTAAGATG 1691
|||||
```

```
551 SerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerLe 567
|||||
1692 TCTCTTGATTTACCTGAGAACAGATGGAAGTCTTTTCTCTCTCTCT 1741
|||||
567 uLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSerS 584
|||||
1742 GTTCCCAACATCCTCTACATCCCTCTCTCTTATTACAAATTCACATGAT 1791
|||||
584 erLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnGlu 600
|||||
1792 CTTTATCACTGAATTCCTCAACCAATATTTCTCTCACTATTGAACGAGAG 1841
|||||
601 SerAlaValLeuAlaThrAlaProArgIleAspAspGluIleProProPr 617
|||||
1842 TCAGCTGTACTAGCAACTGCTCCAAGGATAGATGAATCCCCCTCC 1891
|||||
617 oLeuProValArgThrProGluSerPheIleValValGluGluAlaGlyG 634
|||||
1892 ACTTCTGTACGGACACTGAATCATTTATTGTGTGTTGAGGAAGCTGGAG 1941
|||||
634 luPheSerProAsnValProLysSerLeuSerSerAlaValLysValLys 650
|||||
1942 AATTCTCACCAAAATGTTCCCAAAATCTCTATCTCAGCTGTGAAGTAAAA 1991
|||||
651 IleGlyThrSerLeuGluTrpGlyThrSerGluProLysLysPheAs 667
|||||
1992 ATTGGAACATCACTGGAATGGGTGGACATCTGAACCAAGAAATTTGA 2041
|||||
667 pAspSerValIleLeuArgProSerLysSerValLysLeuArgSerProL 684
|||||
2042 TGACTCTGTGATCTTAGACCAAGCAAGAGTGTAAACTCCGAAAGTCCTA 2091
|||||
584 ysser 585
|||||
2092 AATCA 2096
|||||
seq_name: gb_pat:AR149916
seq_documentation_block:
LOCUS AR149916 3580 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 1 from patent US 6228641.
ACCESSION AR149916
VERSION AR149916.1 GI:15114507
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 3580)
AUTHORS
Jallal B. and Plowman, G.D.
TITLE
Diagnosis and treatment of pTP04 related disorders
JOURNAL
Patent: US 6228641-A 1 08-MAY-2001;
FEATURES
Location/Qualifiers
source
1..3580
/organism="unknown"
BASE COUNT 1213 a 666 c 618 g 1083 t
ORIGIN
alignment_scores:
Quality: 3543.50 Length: 685
Ratio: 5.188 Gaps: 1
Percent Similarity: 99.708 Percent Identity: 99.416
alignment_block:
US-09-600-358A-4 x AR149916 ..
Align seg 1/1 to: AR149916 from: 1 to: 3580
1 MetaspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17
|||||
53 ATGGACCAAGAGAGAAATTCGAGAAAGTTCTGTAAGTTCGAGCCCAAGCAA 102
|||||
17 sLysileThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArgG 34
|||||
```

|||||
103 GAAATTAAGAGGAGCTTTGCCAATGAATTTCTGAAGCTGAAAGGC 152
34 lnsrThrLysTyrLysAlaAspLysThrTyrProThrValAlaGlu 50
153 AATCTACCAAGTACAGGAGAGCAAAACCTATCCTCAACATGTGGCTGAG 202
51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
::: |||||
203 AAGCCCAACAATATCAAGAAACACAGATATAGGATATTTGCCCTATGA 252
67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSert 84
253 TTATAGCCGGTAGAACTATCCCTGATACCTCTGTATGAGGATTCAGCT 302
84 yrIleAsnAlaAsnPheIleLysGlyValTyrClyProLysAlaTyrIle 100
303 ACATCAATGCCAACTTCATTAAGGGAGTTATGGACCCAAAGGCTTATAT 352
101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTyrArgMetil 117
353 GCCACCCAGGGTCTTTATCTACAAACCTCTGGACTTCTGGAGGATGAT 402
117 eTrpGluTyrSerValLeuIleValMetAlaCysMetGluTyrGluM 134
403 TTGGGAATATAGTGTCTTATCATTTGTTATGGCATGCATGGATATGAAA 452
134 etGlyLysLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
453 TGGGAAGAAAGTAGTGTAGCGCTACTGGGCTGAGCCAGAGAGATGCAG 502
151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
503 CTGGAATTTGGCCCTTCTCTGTATCTCTGTGAAGCTGAAAGAGGAATC 552
167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
553 TGATTATATAATCAGGACTCTAAAGTTCAAGTTCAATAGTAGTGAACATCGAA 602
184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSer 200
603 CTATCTACCAAGTTTCATTAAGAATTTGCCAGACCATGATGATACCTTCA 652
201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCystyrGlnGl 217
653 TCTATAGACCTATTCTTGAGCTCATCTGGGATGTACGTTGTTACCAAGA 702
217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
703 GGATGACAGTGTCCCATATGCATTCACATGCAGTGTGGCTGTGGAAGGA 752
234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
753 CTGGTGTATTGTGCTATT...GATTATACATGGATGTGCTTAAAGAT 799
251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267
800 GGGATTAATCCTCAGAACCTTCAGTGTTCAGTTTGATCCGGGAATGCG 849
267 gThrGlnArgProSerLeuValGlnThrGlnGluInTyrGluLeuValT 284
850 GACACAGAGGCCCTTCATTAGTTTCAACCGCAGGAACAATATGAACCTGTCT 899
284 yrAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
900 ACAATGCTGTATTAGAACTATTAAAGAGACAGATGGATGTTATCAGAGAT 949
301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
950 AAACATCTGGAACAGAGAGTCAAGCAAGCATTTGTTATCTCGAGAAAAA 999
317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
|||||

1000 TCACACTCTCCAAGCAGACTCTTATTCTCTCTAATTTACCAAAAAGTACCA 1049
334 hrLysAlaAlaLysMetMetAsnGlnGlnArgThrLysMetGluIleLys 350
1050 CAAAAGCAGCAAAATGATGAACCAACAAGACAAAATGGAATCATAA 1099
351 GluSerSerSerPheAspPheArgThrSerGluIleSerAlaLysGluI 367
1100 GAATCTTCTCTCTTTCACCTTTAGGACTTCTGAAATAAGTGCAAAAGA 1149
367 uLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGluL 384
1150 GCTAGTTTTCACCCCTGCCTAAATCAAGCACTCTCTTTGACTTTCTGGAGC 1199
384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnThr 400
1200 TAAATTAACAGTTTGTACAAAATGCTGCACACCAACATGAATGGCAGACA 1249
401 LysAlaPheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAs 417
1250 AAGGCATTTCCAATAGTTGGGGAGCCTCTTCAGAAGCATCAAAAGTTTGA 1299
417 pLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsnA 434
1300 TTTGGGCTCTCTTTTGTGGGGATGTTCTAATTTCTAAACCTGTAAAT 1349
434 laAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSer 450
1350 CAGCAGGAAGATATTTAATTCAAAGGTGCCAATAACACGGACCAATCA 1399
451 ThrProPheGluLeuIleGlnGlnArgGluThrLysGluValAspSerLy 467
1400 ACTCCTTTTGAATTTGATACAGCAGAGAGAAACCAAGGAGGTGCACAGCA 1449
467 sGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheValG 484
1450 GGAACACTTTTCTTATTTGGAATCTCAACACACATGATTTGTTTGTAG 1499
484 luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
1500 AGATGAGGCTCAAAAAGTAATGCATGTTCTTCAGCAGAACTGAATAT 1549
501 SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnVally 517
1550 TCACCTGCCATATGACTTAACACCAAAATACGTAATGCCCTCAATGTAAA 1599
517 sHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProIleuValG 534
1600 GCACCATGACTTAGTGCTCTTGGTGTATATTCTTACATACCTTTAGTGG 1649
534 luAsnProTyrPheSerSerTrpProProSerGlyThrSerSerLysMet 550
1650 AAAATCCTTATTTTTCATCATGCTCCCAAGTGGTACCAGTTCTAAGATG 1699
551 SerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerle 567
1700 TCTCTTGATTTTACCTGAGAAGCAAGATGGAACGTGTTTTCCTCTCTCT 1749
567 uLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSerS 584
1750 GTTGCCAAACATCTCTACATCCCTCTCTCTTATTACATATCATATGATT 1799
584 erLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnGlu 600
1800 CTTTATCAGTGAATTTCCCAACCAATATTTCTCTACTATTGAACCCAGGAG 1849
601 SerAlaValLeuAlaThrAlaProArgIleAspAspGluIleProProPr 617
1850 TCAGCTGTACTAGCAACTGCTCCAAGGATAGATGATGAAATCCCCCTCC 1899
617 oLeuProValArgThrProGluSerPheIleValValGluGluAlaGlyG 634
1900 ACTTCTGTACGGACACCTGAATCATTTATTGTGGTTGAGGAAGCTGGAG 1949

634 luPheSerProAsnValProLysSerLeuSerSerAlaValLysValLys 650
|||||
1950 AATTCTCACCAGAAATGTTCCCAATCTTATCTCAGCTGTGAAGTAAAA 1999
651 lIleGlyThrSerLeuGluTrpGlyGlyThrSerGluProLysLysPheAs 667
|||||
2000 ATTGGAACATCAGTGGGATGGGTGGAACATCTGAACCAAGAAATTTGA 2049
667 pAspSerValIleLeuArgProSerLysSerValLysLeuArgSerProL 684
|||||
2050 TGACTCTGTGATACTTACACCAAGCAAGAGGTAAACTCCGAAAGTCTTA 2099
684 ySer 685
|||||
2100 AATCA 2104

seq_name: gb_pr:AF150732

seq_documentation_block:

LOCUS AF150732 2877 bp mRNA linear PRI 04-MAY-2000
DEFINITION Homo sapiens protein tyrosine phosphatase (PTP) mRNA, complete cds.

ACCESSION AF150732

VERSION AF150732.1 GI:7688662

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2877)

AUTHORS Huang,C., Zhang,C., Wu,T., Peng,Y., Gu,Y., Zhang,L., Jiang,C.,

Li,X., Han,Z., Wang,X., Chen,Z. and Fu,G.

A novel gene expressed in human adrenal gland

Unpublished

2 (bases 1 to 2877)

AUTHORS Huang,C., Zhang,C., Wu,T., Peng,Y., Gu,Y., Zhang,L., Jiang,C.,

Li,X., Han,Z., Wang,X., Chen,Z. and Fu,G.

Direct Submission

TITLE Submitted (12-MAY-1999) Chinese National Human Genome Center at

Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,

Shanghai 201203, P. R. China

FEATURES

Location/Qualifiers

1..2877

/organism="Homo sapiens"

/db_xref="taxon:9606"

/tissue.type="adrenal gland"

1..2877

/gene="Ptp"

71..2470

/gene="Ptp"

/codon_start=1

/product="protein tyrosine phosphatase"

/protein_id="AAF67472.1"

/db_xref="GI:7688662"

/translation="MDQREILQKFLDEAQSKKITKEEFANEFLLKLRQSTKYKADKTY

/PrtAAEKPNKIKNRYKDTLPYDYSRVELSLITSDSSYINANFKYVGEAYIAT

OQPLSTLLDFRMILWEYSLIIVMACMEYEMKKCEPYWAEPEGMQLFEGFVSVC

EAEKRSYIILRLVKFNSERTIYQHYKNPDHVPSSIDPILDLWVRCQED

DSVPIKCSACGRTGVICAIDYTWMLLKDGIIIPENFSVSLIREMTRQPSLVOTQ

EOYELVNAVELFKRQMDVIRDKHSGTESQAKHCIPENKHTLQADSYSPNLPKSTTK

AAKMMQRTMEIKRESSDFRTSEISAKEDSLAPENQAFCLSEVNSFAPKLPH

HEMPAKAFQIVGSSEASKLNWASLCGMNSNLVCSRKIFNPRCQYDPQINSFELIQ

QENOGDSKPTFLFESQPHDSCFVEMOQAKVMHVSSAELNYSLPYDYSKHQIENASVK

HDSSALGVYSYIPIVENPYRSPWPPSGTSRKMSDLQPEKQDQTFPPSSLLPTSTSL

FSYINSHDSLNSVNTSSIRIEQESAVLATAPRIDDEIPLHFLYGTPESTIVVEAG

ESPNPVKSLGSAVKIGTSLWGGTSEPKKFDSDVILRPSKVLRSKPKSELHQDR

SSPPPLPERTLESFLADECMQASITETSYSPYDMENSTSSKQILKTPGKFTFR

SKSLKILRMKSKSICNSCPPNKPAPESVQSNSSSFLNFGFANRESKPKGPRNPPTWN

I"

BASE COUNT 944 a 583 c 523 g 827 t

ORIGIN

alignment_scores:

Quality: 2977.50 Length: 687

Ratio: 4.660 Gaps: 11

Percent Similarity: 93.013 Percent Identity: 89.229

alignment_block:

US-09-600-358A-4 x AF150732

Align seg 1/1 to: AF150732 from: 1 to: 2877

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17

|||||

71 ATGGACCAAGAGAGAAATCTCAGAAAGTTCTGGATGAGGCCCAAGCAA 120

|||||

17 sLysTleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArg 34

|||||

121 GAAATTAATAAGAGAGAGTTTGGCAATGAATTTCTGAAGCTGAAAGGC 170

|||||

34 InSerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50

|||||

171 AATCTACCAAGTACAGGCGACAGCAAAACCTATCTTACAACCTGGCTGAG 220

|||||

51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67

::: |||||

221 AAGCCCAAGAAATATCAGAAAAACAGATATAAGGATATTTTCCCTATCA 270

|||||

67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSer 84

|||||

271 TTATAGCGGGGTAGAACTATCCCTGATAACCTCTGATGAGGATTCAGCT 320

|||||

84 YrIleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100

|||||

321 ACATCAATGCCAACTTCATTAAGGAGTTTATGGACCGAGGCTTATATT 370

|||||

101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetIl 117

|||||

371 GCCACCCAGGGTCCTTTATCTACAACCCCTCTGGACTTCTGGAGGATGAT 420

|||||

117 eTrpGluTyrSerValLeuIleIleValMetAlaCysMetGluTyrGluM 134

|||||

421 TTGGGAATATAGTGTCCTTATCATTTGTTATGGCATGGAGTATGAAA 470

|||||

134 etGlyLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150

|||||

471 TGGGAAGAAAAGTGTGAGCGCTACTGGCTGAGCCAGAGAGATGCGAG 520

|||||

151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167

|||||

521 CTGGAATTTGGCCCTTTCTCTGTATCTCTGTAAGCTGAAAGTGAAGAAATC 570

|||||

167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184

|||||

571 TGATTATATAATCAGGACTCTAAAAGTTAAGTTCAATAGTGAACCTCGAA 620

|||||

184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSer 200

|||||

621 CTATCTACCACTTTCATTACAGAATTTGGCCAGACCATGATGTACCTTCA 670

|||||

201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnGl 217

|||||

671 TCTATAGACCCCTATTCTTGGAGCTCATCTGGGATGTCCTGTTACCCAAGA 720

|||||

217 uAspAspSerValProIleCysIleIleHisCysSerAlaGlyCysGlyArgT 234

|||||

721 GGATGACAGTGTTCCCATATCATTCATCTGAGTGTGCTGCTGTGGAGGA 770

|||||

234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250

|||||

771 CTGGTGTATTGTGCTATT...GATTATACATGGATGTTGCTAAAAAGAT 817

|||||

251 GlyIleLeuProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267

|||||

818 GGGATAATTCCTGAGAACTTCAGTGTGTTTTCAGTTTGATCCGGGAAATCGC 867

```
267 gthrGlnArgProSerLeuValGlnThrGlnGlnGlnTyrGluLeuVal 284
|||||
868 GACACAGAGCCCTTCATTAGTTCACACGAGGAGCAATATGACTGGTCT 917
284 yrhAnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
|||||
918 ACAATGCTGTATTAGAACTATTTAAGAGACAGATGGATGTTATCAGAGAT 967
301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
|||||
968 AAACATCTCGGAACAGAGACTCAAGCAAGCAATGTATTCTCGAGAAAA 1017
317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
|||||
1018 TCACACTCTCAAGCAGACTCTATTCTCTTAATTTACCAAAAGTACCA 1067
334 hrLysAlaLysMetMetAsnGlnGlnArgThrLysMetGluIleLys 350
|||||
1068 CAAAAGCAGCAAAATGATGAACCAACAAAGGACAAAATGGAATCAAA 1117
351 GluSerSerPheAspPheArgThrSerGluIleSerAlaLysGlu 367
|||||
1118 GAATCTCTCTCTTACCTTAGGACTTCTGAAATAGTCAAAAGAA 1167
367 uLeuValLeuHisProAlaLysSerThrSerPheAspPheLeuGlu 384
|||||
1168 CGTAAGTTGGCCCG...GAAACACAGGCTTTTTTGGCTTCGGAAG 1214
384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnThr 400
|||||
1215 TAAATCCAGTTTTGCAAAATGTTG...CCCCACCATGAATGCCAGCA 1261
401 LysAlaPheProIleValGlyGluProLeuGlnLysHisGlnSerLeu 417
|||||
1262 AAGGCTTTCAAATAGTGGGA...AGCTCTTCAGAAGCATCAAGT...T 1305
417 sLeuGlySerLeuLeuGluGlyCysSerAsnSerLysProValAsn 433
|||||
1306 GAATGGGCTCTCTTGTGTAGGATGCTA...ATTCTAACCTTGAT 1352
434 AlaAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSe 450
|||||
1353 GCAGCAGAAAGATATTATATCCAAAGTGCCAAT...ACCGGACCAAAATC 1399
450 rThrProPheGluLeuIleGlnGlnArgGluThrLysGluValAspSer 467
|||||
1400 AACTCT.TTTGAATTGATCCAGCAA...GAGAACCAAGGAGGTGACAGCA 1445
467 yGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheVal 483
|||||
1446 AGAAACTTCTTATT...GAATCTCAACCATGATCTTGTGTTGTA 1492
484 GluMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTy 500
|||||
1493 GAGATGCGAGCTCAAAAGTAATGATGCTTCTTCAGCAGAACTGAATTA 1542
500 rSerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnVal 517
|||||
1543 TTCACCTGCCATATGACTCTAAACACCAAAATACGTAATGCTCTAATGTAA 1592
517 yHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuVal 533
|||||
1593 AGCACCATGACTCTAGTCTCTGGGTATATTCTTACATACCTTTAGTC 1642
534 GluAsnProTyrPheSerSerTrpProProSerGlyThrSerSerLysMe 550
|||||
1643 GAAATCCTTATTTTTCATCATGCGCTCCCAAGTGGTACCAAGTTCTAAGAT 1692
550 tSerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerL 567
|||||
1693 GTCTCTGTATTACCTCAGAGAAGATGGAAGTGTGTTTCTCTCTCTCTC 1742
```

```
567 euLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSer 583
|||||
1743 TGTGTGCAACATCCTCTACATCCTCTCTCTTATTACAAATTACATGAT 1792
584 SerLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnG 600
|||||
1793 TCTTTATCATCTGAATCTCCACCAATATTTCTCTCAGTATTGAACAGGA 1842
600 uSerAlaValLeuAlaThrAlaProGlyIleAspAspGluIleProProp 617
|||||
1843 GTCAGCTGTACTAGCAACTGCTCCAGGATAGATGATGAAT.CCCCTC 1891
617 roLeuProValArg.ThrProGluSerPheIleValValGluGluAlaG 633
|||||
1892 CACTTCTGTACGGGACACCTGAATCATTTATTGTGTTTTCAGGAAGTGG 1941
633 yGluPheSerProAsnValProLysSerLeuSerSerAlaValLysVal 650
|||||
1942 AGAATTCTCACCAATGTTCCCAATATCTTATCTCAGCTGTGAAGTAA 1991
650 ysIleGlyThrSerLeuGluTrpGlyGlyThrSerGluProLysLysPhe 666
|||||
1992 AAATTGGACATCCTCGGAATGGGTGGAACTCTGAACCAAGAAATTT 2041
667 AspAspSerValIleLeuArgProSerLysSerValLysLeuArgSerPr 683
|||||
2042 GATGACTCTGTACTTAGACCAAGCAAGAGTAAACTCCGAGTCC 2091
683 oLysSer 685
|||||
2092 TAAATCA 2098
```

seq_name: gb_ro:MUSPROTYPH

seq_documentation_block:

LOCUS MUSPROTYPH 2734 bp mRNA linear ROD 27-APR-1993

DEFINITION Mouse protein tyrosine phosphatase (70zpep) mRNA, complete cds.

ACCESSION M90388

VERSION 1 GI:200522

KEYWORDS protein-tyrosine phosphatase.

SOURCE Mus musculus cDNA to mRNA.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2734)

AUTHORS Matthews,R.J., Bowne,D.B., Flores,E. and Thomas,M.L.

TITLE Characterization of hematopoietic intracellular protein tyrosine

phosphatases: Description of a phosphatase containing an SH2 domain

and another enriched in proline-, glutamic acid-, serine-, and

threonine-rich sequences

Mol. Cell. Biol. 12, 2396-2405 (1992)

JOURNAL 92236615 Location/Qualifiers

MEDLINE 1..2734

FEATURES /organism="Mus musculus"

source /db_xref="taxon:10090"

1..2734

gene /gene="70zpep"

CDS 113..2521

/gene="70zpep"

/EC_number="2.7.1.112"

/codon_start=1

/product="protein tyrosine phosphatase"

/protein_id="AAA3994.1"

/db_xref="GI:200523"

/translation="MDQREILQQLKEAKKKLNSEEFASEFLKLRKSTKTKYADKIY

PTTVAQPKNKIKKRYKIDLPYDHSLSLSLSDSSYINASIKIGYIGPKYAIAT

OGPLSTLLDFWRMIWEYRILVIVMACMEFEMKKCKERYWAEPEGTQLQFGPFSIC

EAEKKKSDYKIRTLAKENNETRIIYOFHYKNWPDVPSISDIPILQLIWMRCYQED

DCVPICHCAGCGRTGICAVDVTWMLLKDGIIPKNSVFNLIQEMRTQRESLQTO

EQYELVYSAVLELFRHMDVISDNHLGREIQKQCSIPQSLTVEADSCPLDLPKNAMR

DYKTNQHSKQGAEEASTGSLGRTSTMNAEEELVLHSAKSSPSFNCLELNCCGN

KAVITRNQARASPVVYGEPLQYQSLDFGSMFLGSCPSALPINTADRVHNSGPKVRT

KSTPFLIOQRKINDIADVGDGFSCLSEQLHEHYSLRELQVQRVAHVSSBELNYSLPGA
CDASCVPRIHSPGALRVHLVYLSLAEDPYFSSSPNSADSKMSFDLPEKODGATSPGALL
PASSTSFYVSNPHDSLWNTLTSFSPPLNOETAVEAPSRRTDDEIPPLPRTPESE
IVVEEAGEPSRVTESLPLVTFGASPEGSTSEKSHDSVGFETPSKNVYKLRSPKSDR
HODGSPPLPRTLESFLAEDCQAOAVQTSSTPYPETENSTSSKQTLRTPKRS
FTKSKLKIFRNKMSVCSNSSPSKPTERVQPRKNSSSLNFGFGRFSPKPGPRNPPS
ANW*

BASE COUNT 787 a 629 c 601 g 717 t
ORIGIN

alignment_scores:

Quality: 2356.50 Length: 692
Ratio: 4.028 Gaps: 6
Percent Similarity: 84.538 Percent Identity: 68.497

alignment_block:

US-09-600-358A-4 x MUSPROTYPH ..

Align seg 1/1 to: MUSPROTYPH from: 1 to: 2734

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17
|||||
113 ATGGACCAAGAGAAATCTCAGCACTACTGAAAGAGGCCGCAAAAA 162
17 sLysileThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArgG 34
|||||
163 GAAACTCAACAGTCAGGAGTTTGCAGTGAATTTCTGNAAGCTGAAAGGC 212
34 lnSerThrLysTyLysAlaAspLysThrTyProThrThrValAlaGlu 50
|||||
213 AATCAGCAAGTACAGGCCGACAAAATCTATCTACAACTGTGCTCAG 262
51 AsnAlaLysAsnIleLysLysAsnArgTyLysAspIleLeuProTyAs 67
::: |||||
263 AGGCCCAAGAAATATCAAGAAAAACAGATACAAAGGATATTTGCCCTATGA 312
67 pTySerArgValGluLeuSerLeuIleThrSerAspGluAspSerSert 84
:|||||
313 TCACAGCTGTGACAGCTGCTCTGTAACTTCGATGAGGATCCAGTT 362
84 yrIleAsnAlaAspPheIleLysGlyValTyGlyProLysAlaTyIle 100
:|||||
363 ATATCAATGCCAGCTTTTAAAGGCTGTCTATGACCAAGCTTATATT 412
101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetIl 117
|||||
413 GCTACCCAGGCTCTTTATCTACAACTCTCCTGGACTTCTGGAGGATGAT 462
117 eTrpGluTySerValLeuIleValMetAlaCysMetGluTyGluM 134
|||||
463 TTGGAGTACCGCATCTTGGTCATGTGTCATGGCATGTATGGAGTTGAAA 512
134 etGlyLysLysCysGluArgTyTrpAlaGluProGlyGluMetGln 150
|||||
513 TGGGAAAAAATAATGTGAGCGTTATTGGGCCGCAACAGGAGAAACGCG 562
151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
|||||
563 CTGCAATTTGGCCCTTTCTTATATCTGTGAAGCTGAGAAAAAGAAATC 612
167 rAspTyIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
|||||
613 TGATATAAATACAGGACTCTGAAGGCCAAGTTCAATAATGAAGAACTGAA 662
184 hrIleTyGlnPheHisTyLysAsnTrpProAspHisAspValProSer 200
|||||
663 TTAUTTACCACTTTCAUTTATAAAGATTGGCCAGACCATGATGTGCCITCA 712
201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyGlnGl 217
|||||
713 TCTATAGACCCCTATTCTCAGCTCATCTGGGATATGCGTTGTTACCAAGA 762

217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
|||||
763 AGATGACTGTCTCTCTATATCATCTCACTGTCAGTCCGCGCTCGGGAAGA 812
234 hrGlyValIleCysAlaIleValAspTyThrTrpMetLeuLeuLysAsp 250
|||||
813 CAGTGTCTATTGTGCT..GTTGATTATACATGATGCTGCTGAAAGAT 859
251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267
|||||
860 GGGATAATCTCCTAAGACCTCAGTGTGTTTTAATTTGATTCAGGAGATCG 909
267 gThrGlnArgProSerLeuValGlnThrGlnGluGlnTyGluLeuValT 284
|||||
910 AACACAGAGGCTTCGCTAGTTCAAACTCAGGAACAGTACGAACTGGTCT 959
284 yrAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgasp 300
|||||
960 ACAGTGTCTGTATTAGAGCTGTTAAGAGGCACATGGATGTTATCTCTGAT 1009
301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
::: |||||
1010 AATCACCTTGAAGAGAGATTCAGACACATGCTCAATTCCTGAAACAAG 1059
317 nHisThrLeuGlnAlaAspSerTySerProAsnLeuProLysSerThr 334
: |||||
1060 CCTCAGCTGACAGCTGACTCTGCTCTCTGATTTACCAAAAACGCCA 1109
334 hrLysAlaAlaLysMetMetAsnGln.....GlnArgThrLysMet 347
::: |||||
1110 TGAGGGATGTGAAAACGACAAACAGCAGCAGTACAAACAGGGGCTGAAGCG 1159
348 GluIleLysGluSerSerSerPheAspPheArgThrSerGluIleSerAl 364
|||||
1160 GAGAGCTGAGGCTCTCTCTGGCCTTAGGACTTCTACGATGAATGC 1209
364 aLysGluGluLeuValLeuHisProAlaLysSerSerThrSerPheAsp 381
:|||||
1210 CGAGGAAGATTGTTTGCACCTCGCTAAATCAAGCCCTCTTTTAAT 1259
381 heLeuGluLeuAsnTySerPheAspLysAsnAlaAspThrThrMetLys 397
|||||
1260 GTTTAGAGCTAAACTGGCGGTGTAAACAAAGGCTGTATCAACCCAGGAAC 1309
398 TrpGlnThrLysAlaPheProIleValGlyGluProLeuGlnLysHisGl 414
|||||
1310 GGGCAGCAAGGCTTCTCCAGTCTGGGAGAGCCCTTCAGAAAGATCA 1359
414 nSerLeuAspLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysp 431
|||||
1360 AAGCTGGAATTTGTTCCATGTTTGGGTCTGTCGTCTAGTCTCTGTC 1409
431 roValAsnAlaAlaGlyArgTyThrPheAsnSerLysValProIleThrArg 447
|||||
1410 CCATAAACACACAGCGACAGGTATCACATTCATAAGGGCGCGTAAACGG 1459
448 ThrLysSerThrProPheGluLeuIleGlnArgGluThrLysGluVa 464
|||||
1460 ACCAACTCACTCCCTTTGAATGATTCAGCAGAGAAACAAATGACTT 1509
464 lAspSerLysGluAsnPheSerTyLeuLeuSerGlnProHisAspSerC 481
: |||||
1510 GGCCGTGGGACAGCGTTTTCATGCTGGAATCTCAGCTGCATGAGCAT 1559
481 ysPheVal...GluMetGlnAlaGlnLysValMetHisValSerSerAla 496
::: |||||
1560 ACAGTCTCAGGAGCTGCAGGTGCAAGAGTGGCCCATGTTTCTTCAGAA 1609
497 GluLeuAsnTySerLeuProTyAspSerLysHisGlnIleArgAsnAl 513
|||||
1610 GAGCTGAATTTATCTACCTGCCT.....GGTGCCTGTGATGC 1644
513 aserAsnValLysHisHisAspSerSerAlaLeuGlyValTySerTyI 530

1645 GTCTGTGTGCCCCGGCACCCCGCGCTTTGAGAGTGCATCTGTACA 1694
 530 leProLeuValIleuAsnProTyrPheSerSerTrpProSerGlyThr 546
 1695 CATCTTTAGCGGAAGATCCTTATTTTTCATCATCCCTCCGAATAGTCT 1744
 547 SerSerLysMetSerLeuAspLeuProGluLysGlnAspGlyThrValph 563
 1745 GATTCGAAGATGCTTTTATGATCCCTCGAGAAACAGATGGAGCCATTC 1794
 563 eProSerSerLeuLeuProThrSerSerThrSerLeuPheSerTyrTyra 580
 1795 CCGTGGCGCTCTATTGCCAGCCTCTTCACAACTCCTCTTTATATGACA 1844
 580 snSerHISerSerLeuSerLeuAsnSerProThrAsnIleSerSerLeu 596
 1845 ACCCACAGCAGCTCCCTAGTGATGAACACTCTGACGACGCTTTTCCGCCACG 1894
 597 LeuAsnGlnGluSerAlaValIleuAlaThrAlaProArgIleAspAspG 613
 1895 TTAACCAAGAGACAGCTGTAGAGCTCTCTCGGAGGACAGATGATGA 1944
 613 uLeProProLeuProValAla-gThrProGluSerPheIleValValG 630
 1945 AATCCCCCGGCACCTCCCTGAACGGACACCCGAGCTCTTTATTGTGGTTG 1994
 630 luGluAlaGlyGluPheSerProAsnValProLysSerLeuSerSerAla 646
 1995 AGGAAGCCGGAGAGCCCTCACACGCTGTACCGAATCCTTA.....CCT 2038
 647 valLysValLysIleGlyThrSerLeuGluTrpGlyGlyThrSerGluPr 663
 2039 CTGTGGTGAACATTGTGGAGCATCACCAAGATGCAGTGGGACATCTGAAT 2088
 663 oLysLysPheAspAspSerValIleLeuAlaProSerLysSerValLysL 680
 2089 GAAGAGC...CATGACTCTGTAGGTTTACACCAAGCAAGATGTGAAC 2135
 680 euArgSerProLysSerGlyLysAsn 688
 2136 TCCGAAGTCCCAAAATTCAGATCGACAT 2161

as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/C.elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1> RP5-107303 is from the library RPCI-5 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pCYPAC2

FEATURES
source

```

true light end of chromosome 11p24.2 is at 119109 in this sequence.
RES
source
Location/Qualifiers
1. .143618
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="p13.1-13.3"
/clones="RF5-107303"
/clone_lib="RPC1-5"
257. .322
/note="MER5A repeat: matches 41. .109 of consensus"
763. .1053
/note="AluSx repeat: matches 1. .312 of consensus"
1348. .1698
/note="L2 repeat: matches 2070. .2417 of consensus"
1919. .1969
/note="L2 repeat: matches 2655. .2705 of consensus"
2305. .2446
/note="MIR repeat: matches 40. .176 of consensus"
5652. .5927
/note="L1ME repeat: matches 5524. .5813 of consensus"
6295. .6428
/note="BC200 repeat: matches 2. .147 of consensus"
6653. .6700
/note="8 copies 6 mer acacac 97% conserved"
7925. .7980
/note="28 copies 2 mer aa 75% conserved"
8071. .8369
/note="AluSx repeat: matches 1. .296 of consensus"
9048. .9342
/note="AluSx repeat: matches 1. .298 of consensus"
9793. .10359
/note="L1MC1 repeat: matches 5526. .6320 of consensus"
10574. .10730
/note="MIR repeat: matches 95. .262 of consensus"
10759. .10907
/note="MIR repeat: matches 2. .139 of consensus"
10908. .11219
/note="AluJb repeat: matches 1. .296 of consensus"
11220. .11315
/note="MIR repeat: matches 139. .237 of consensus"
12210. .12464
/note="AluSg repeat: matches 1. .291 of consensus"
12916. .13149
/note="MER58B repeat: matches 70. .341 of consensus"
13150. .13435
/note="AluSx repeat: matches 1. .298 of consensus"
13458. .13805
/note="L1MD2 repeat: matches 4851. .5213 of consensus"
13806. .14115
/note="AluJb repeat: matches 1. .299 of consensus"
14118. .14430
/note="AluSx repeat: matches 1. .312 of consensus"
14431. .14563
/note="L1MD2 repeat: matches 5213. .5314 of consensus"

```



```
repeat_region 14565..14859
/notes="AluSg repeat: matches 1. .294 of consensus"
repeat_region 14861..14888
/notes="14 copies 2 mer ta 92% conserved"
repeat_region 14910..15215
/notes="L1MD2 repeat: matches 5302. .5609 of consensus"
repeat_region 15230..15560
/notes="MER2 repeat: matches 2. .345 of consensus"
repeat_region 15561..16146
/notes="L1MD2 repeat: matches 5643. .6210 of consensus"
repeat_region 16270..16479
/notes="MIR repeat: matches 48. .262 of consensus"
repeat_region 16840..17032
/notes="MER3 repeat: matches 2. .207 of consensus"
repeat_region 17472..17501
/notes="15 copies 2 mer ac 100% conserved"
misc_feature 1839..19765
/notes="CpG Island"
/evidence=not_experimental
20385..20649
/notes="AluJo repeat: matches 1. .297 of consensus"
repeat_region 20642..20734
/notes="FLAM_A repeat: matches 4. .116 of consensus"
repeat_region 22760..23065
/notes="AluSx repeat: matches 1. .306 of consensus"
repeat_region 23275..23506
/notes="AluJb repeat: matches 1. .225 of consensus"
repeat_region 23540..23584
/notes="L2 repeat: matches 2647. .2694 of consensus"
repeat_region 23590..23671
/notes="AluJ/FLAM repeat: matches 1. .83 of consensus"
repeat_region 23714..24013
/notes="AluSx repeat: matches 1. .297 of consensus"
repeat_region 24532..24817
/notes="AluJb repeat: matches 17. .293 of consensus"
repeat_region 25033..25166
/notes="FLAM_C repeat: matches 11. .143 of consensus"
repeat_region 25261..25549
/notes="AluSx repeat: matches 1. .289 of consensus"
repeat_region 25675..25877
/notes="MLTJ repeat: matches 2. .214 of consensus"
repeat_region 25998..26083
/notes="Charliel repeat: matches 3. .92 of consensus"
repeat_region 27261..27556
/notes="AluSx repeat: matches 3. .289 of consensus"
repeat_region 28123..28175
/notes="L2 repeat: matches 2678. .2733 of consensus"
repeat_region 28395..28708
/notes="AluSg repeat: matches 1. .310 of consensus"
repeat_region 29346..29766
/notes="L2 repeat: matches 2085. .2448 of consensus"
repeat_region 29767..30052
/notes="AluSg repeat: matches 1. .287 of consensus"
repeat_region 30053..30345
/notes="L2 repeat: matches 2448. .2750 of consensus"
repeat_region 31191..31486
/notes="AluSg repeat: matches 1. .296 of consensus"
repeat_region 31804..31932
/notes="AluJb repeat: matches 1. .127 of consensus"
repeat_region 32054..32075
/notes="11 copies 2 mer aa 100% conserved"
repeat_region 32503..32780
/notes="AluJo repeat: matches 1. .288 of consensus"
repeat_region 32791..32817
/notes="9 copies 3 mer tgt 96% conserved"
repeat_region 32817..33099
/notes="AluSg repeat: matches 1. .282 of consensus"
repeat_region 33156..33450
/notes="AluJb repeat: matches 6. .301 of consensus"
repeat_region 33747..34055
/notes="AluSx repeat: matches 1. .309 of consensus"
repeat_region 34584..34775
/notes="L1ME repeat: matches 5644. .5823 of consensus"
```

```
repeat_region 34786..34930
/notes="FLAM_C repeat: matches 12. .142 of consensus"
repeat_region 34931..35225
/notes="L1ME repeat: matches 5265. .5563 of consensus"
repeat_region 35257..35727
/notes="L1ME repeat: matches 4762. .5233 of consensus"
repeat_region 35743..36021
/notes="AluJb repeat: matches 1. .287 of consensus"
repeat_region 36340..36488
/notes="L1PA15 repeat: matches 6009. .6157 of consensus"
repeat_region 37329..37776
/notes="L1MC5 repeat: matches 7275. .7680 of consensus"
repeat_region 38067..38121
/notes="L2 repeat: matches 2681. .2735 of consensus"
repeat_region 38346..38564
/notes="L1MB5 repeat: matches 5921. .6142 of consensus"
repeat_region 39036..39268
/notes="MIR repeat: matches 20. .262 of consensus"
repeat_region 39278..39564
/notes="AluSx repeat: matches 1. .294 of consensus"
repeat_region 39575..39657
/notes="AluJo/FRAM repeat: matches 206. .289 of consensus"
repeat_region 41653..41957
/notes="AluJb repeat: matches 5. .293 of consensus"
repeat_region 42324..42634
/notes="AluY repeat: matches 1. .311 of consensus"
repeat_region 43028..43056
```

```
alignment_scores:
  Quality: 1490.50      Length: 345
  Ratio: 4.985          Gaps: 2
  Percent Similarity: 86.667  Percent Identity: 86.087

alignment_block:
US-09-600-358A-4 x AL137856/rev ..
```

```
Align seg 1/1 to reverse of: AL137856 from: 1 to: 143618

306 GluSerGlnAlaLysHisCysIleProGluLysAsnHisThrLeuGlnAl 322
:::|||||
45710 AAGAGTCAAGCAAGCATGTATTCTCGAGAAAATCACACTCTCCAAGC 45661
322 aASpSerTyrSerProAsnLeuPro.Lys..... 331
|||||
45660 AGACTCTATTCTCCTTAATTTACCAAAAGGTGAGTACATAATTTATATAT 45611
331 ..... 331

45610 CTAAGTTGACTAATTGTGTCTGGAGATGAATCTCATGCTTATACATC 45561
331 ..... 331

45560 ATTGTGATGCTAGGAAGCGGCTGTTGGGTATGGTGTTCATCATTCAGC 45511
332 .....SerThrLysAlaAlaLysMetMetAsnGlnGln 343
|||||
45510 TTTCTTTTGTGTTAGTACCACAAAGCAGCAAAATGATGACCAACAA 45461
344 ArgThrLysMetGluLeuLysGluSerSerPheAspPheArgThrSe 360
|||||
45460 AGGACAAAATGGAATCAAGAAATCTTCTTCCTTTGACTTTAGGACTTC 45411
360 rGluLeuSerAlaLysGluLeuValLeuHisProAlaLysSerSerT 377
|||||
45410 TGAATAAGTGCAAAAGAGAGTAGTTTGCACCCCTGCTAAATCAAGCA 45361
377 hrSerPheAspPheLeuGluLeuAsnTyrSerPheAspLysAsnAlaAsp 393
|||||
45360 CTTCCTTTTGACTTCTCGAGCTAAATTTACAGTTTGGACAAAATGCTGAC 45311
394 ThrThrMetLysTrpGlnThrLysAlaPheProIleValGlyGluProLe 410
|||||
45310 ACAACCATGAATGGCAGACAAAGGCATTTCCCAATAGTTGGGGAGCCTCT 45261
```



```
80 luAspSerSerTyrIleAsnAlaAsnPheIleLysGlyValTyrGlyPro 96
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
270 AAGATTTCAGACTATATCAATGCAAAATTTATAAGGGCGCTCTATGGCCA 319
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
97 LysAlaTyrIleAlaThrGlnGlyProLeuSerThrThrLeuLeuAspPh 113
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
320 AAAGCATATGTAGCAACTCAAGGACCTTTAGCAAAATACAGTAATAGATT 369
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
113 eTripArgMetIleTripGluTyrSerValLeuIleValMetAlaCysM 130
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
370 TTGGAGATGGTATGGAGATATAATGTGTGCATCATTTGTAATGGCCGTGC 419
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
130 etGluTyrGluMetGlyLysLysCysGluArgTyrTriAlaGluPro 146
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
420 GAGAAATTTGAGATGGGAAGGAAATGAGGCGCTATTGGCCCTTTGTAT 469
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
147 GlyGluMetGlnLeuGluPheGlyProPheSerValSerCysGluAlaG 163
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
470 GGAGAAGACCCCATACGTTTGGACCAATTTAAATTTCTTGTGAGGATGA 519
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
163 uLysArgLysSerAspTyrIleAlaGthrThrLeuLysValLysPheAsn 180
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
520 ACAGACGACAGACACTCTCATCAGACACTCTTACTTGAATTCAAA 569
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
180 erGluThrArgThrIleTyrGlnPheHisTyrLysAsnTriProAspHis 196
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
570 ATGAATCTCTAGGCTCTATCAGTTTCATTATGTGAACCTGGCCAGCCAT 619
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
197 AspValProSerSerIleAspProIleLeuGluLeuIletrpaspValar 213
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
620 GATGTTCTCTTCATCATTTGATTCTATTCTGCACATGATAAGCTTAATG 669
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
213 qCysTyrGlnGluAspSerValProIleCysIleHisCysSerAlaG 230
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
670 GAAATATCAAGACATGAAGATGTCTCTATTGTATTTCATTCGACGTGAG 719
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
230 lyCysGlyArgThrGlyValIleCysAlaIleValAspTyrThrTrpMet 246
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
720 GCTGTGGAAGAACAGGTGGCCATTGTGCCATA...GATTATACGTGGAAT 766
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
247 LeuLeuLysAspGlyIleIleProGluAsnPheSerValPheSerLeuI 263
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
767 TTACTAAAGCTGGGAAATACCAGAGGAATTAATGATTATTAATTAAT 816
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
263 eArgGluMetArgThrGlnArgProSerLeuValGlnThrGlnGluGln 280
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
817 ACAGAAATGAGACACACAGGAGGATTCGCAGTACAAACAAAGGAGCAAT 866
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
280 yrGluLeuValTyrAsnAlaValLeuGluLeuPheLysArgGlnMetAsp 296
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
867 ATGAACCTGTTTCATAGAGCTATTGCCCACTGTTTGAAACACAGCTACAA 916
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
297 Val.....IleArgAsp.....LysHisSe 303
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
917 CTATATGAAATTCATGGAGCTCAGAAAATTCCTGATGCAGTGAATGAAAT 966
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
303 rGlyThrGluSerGlnAlaLysHisCysIleProGluLysAsnHisrL 320
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
967 TACACTGAACATAGTCAGCTCCATAGAGCCTGAAAAACAA..... 1009
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
320 euGlnAlaAspSerTyrSerProAsnLeuProLysSerThrThrLysAla 336
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1010 .....GATTCTCTCTCCAAAACCCACCAAGGACCCGCACT..... 1045
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
337 AlaLysMetMetAsnGlnArgThrLysMetGluIleLysGlu..... 351
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1046 .....TGCCCTTGTGAAGGGGATGCTTAAAGAGAAATACTGCAGCCACC 1089
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
352 .....SerSerSerP 355
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1090 GGAACTCATCCAGTGCACCCACCTTCAGACACCTTCCTCCCTTCAGCTT 1139
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
355 heAspPheArgThrSerGluIleSerAlaLysGluGluLeu..... 368
```

```
1140 TTCCAACAGTCACCTACTCTGTGGCAGGACAATGATAGATACCATCCAAG 1189
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
369 ...ValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGluLe 384
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1190 CCAGTGTTCATATGGTTTCATCAGAACACATTCAGCAGACCTCAACAG 1239
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
384 uAsnTyrSer..... 387
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1240 AAACATATAGTAAATCAACAGAACTCCAGGGAAAAATGAATCAACAATTG 1289
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
388 .....PheAspLysAsnAlaAspThrThrMetLysTrpGlnThrLysAla 402
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1290 AACAGATAGATAAAAAATTTGACAGCAAAATTTAAGTTTGTAGATTAAAG 1339
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
403 PheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAspLeuGl 419
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1340 GTCCCTCTCCAAGAGGACCA.....AAAAGTTTGTATGGGAA 1377
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
419 ySerLeuLeuPheGluGly.....C 426
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1378 CACACTTTGAATAGGGGACATGCAATTAATAATTAATCTGCTTCACCTT 1427
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
426 ys.....SerAsnSerLysPro.....ValAsn 433
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1428 GTATAGCTGATAAAATCTTAAGCCACAGGAATTAAGTTCAGATCTAAAT 1477
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
434 AlaAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSe 450
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1478 GTCGCTGATACTTCCCAAGAAATCTTGTGTGACCTGCAGTGTAAACACAATC 1527
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
450 r.....ThrProPheGluIleGlnGlnArgGluPhrL 462
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1528 AAACAAAGTTTCAGTTACTCCACCAGAGAATCCAGAAATTCACACACAC 1577
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
462 ys.....GluValAspSerLysGluAsnPheSerTyr 472
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1578 CTCCAAGGCCAGACCGCTTGCTCTTGATGAGAAAGGACATGTAACGTGG 1627
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
473 LeuGluSerGlnProHisAspSerCysPheValGluMetGlnAlaGlnLy 489
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1628 TCATTTCATGGACCTGAAATGCC.....ATACCCATACCTGA 1665
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
489 sValMetHisValSerSerAlaGluLeuAsnTyr..... 500
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1666 TTTATCTGAAGGCAATTCCTCAGATATCACTATCAAACTAGGAAAACGTG 1715
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
501 ..SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnVal 516
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1716 TGACTTTAAACACCAAGTCCTACACACAAGTTGAAACACCTGATCTTGTG 1765
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
517 LysHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuVa 533
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1766 GATCATGATAACACTTCA.....CCACTCT 1791
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
533 lGluAsnProTyrPheSerSerTrpProSerGlyThrSerSerLysM 550
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1792 CAGAACACCCCTCAGTTTACTAATCCACTCACTCTGATGACTCA.... 1837
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
550 etSerLeuAspLeuProGluLys...GlnAspGlyThrValPheProSer 565
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1838 .....GACTCAGATGAAAGAAACTCTGATGCTGTGCTGACCCAGAAT 1879
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
566 SerLeuLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHi 582
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1880 AAAACTAATATTTCAACAGCAAGTGCACACAGTTTCTGCTGCCACTAGTAC 1929
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
582 sSerSerLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnG 599
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1930 TGAAGCATTTCTACTAGGAAAGTATGGCCAAATGTCCATTCCTGCTAGACATA 1979
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
599 lnGluSerAlaValLeuAlaThrAlaProArg.....IleAsp 611
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```



```

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
420 GAGAAATTTGAGATGGGAAGGAAAAAATGTGAGCGCTATTGGCCTTTGTAT 469
147 GlyGluMetGlnLeuGluPheGlyProPheSerValSerCysGluAlaG1 163
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
470 GGAGAAGACCCCAACATCAACCTTGGCAACATTTAAATTTCTTGTGTGAGGATGA 519
163 uLysArgLysSerAspTyrIleAlaArgThrLeuLysValLysPheAsnS 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
520 ACAGCAAGACAGACACTACTATCATCAGACACTCTTACTTTGAATTTCAA 569
180 erGluThrArgThrIleTyrGlnPheHisTyrLysAsnTrpProAspHis 196
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
570 ATGAATCTCGTAGGCTGTATCAGTTTCATTATGTGAACCTGGCCAGACCAT 619
197 AspValProSerSerIleAspProIleLeuGluLeuIleTrpAspValar 213
620 GATGTCTCTCATCATTTGATTTCTATTCTTGACATGATAAGCTTAATGAG 669
213 gCysTyrGlnGluAspSerValProIleCysIleHisCysSerAlaG 230
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
670 GAAATATCAAGACATGAAGATGTTCTTATTTGTTATTCATTGCGATGCGAG 719
230 lyCysGlyArgThrGlyValIleCysAlaIleValAspTyrThrTrpMet 246
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
720 GCTGTGGAAGAACAGGTGCCATTTGTGCCATA...GATTATACGTGGAAT 766
247 LeuLeuLysAspGlyIleIleProGluAsnPheSerValPheSerLeu1 263
767 TTACTAAAGCTGGGAAATACCGAGAGATTTAATGTTATTTAATTTAAT 816
263 eArgGluMetArgThrGlnArgProSerLeuValGlnThrGlnGluGlnT 280
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
817 ACAAGAAATGAGACACACAAAGCGCTTCGAGTACAAACAAGAGGAGCAAT 866
280 yrGluLeuValTyrAsnAlaValLeuGluLeuPheLysArgGlnMetAsp 296
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
867 ATGAACCTGTTTCATAGAGCTATTGCCAACCTGTTTGAAACACAGCTACAA 916
297 Val.....IleArgasp.....LysHisSe 303
917 CTATATGAATTCATGAGCTCAGAAAAATTCCTGATGAGGTGAATGAAT 966
303 rGlyThrGluSerGlnAlaLysHisCysIleProGluLysAsnHisThrL 320
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
967 TAACACTGAAACATGTCACCTCCATAGACCTCGAAAAACAA..... 1009
320 euGlnAlaAspSerTyrSerProAsnLeuProLysSerThrThrLysAla 336
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1010 .....GATTCCTCTCTCCAAACACCAAGGACCCGCGAGT..... 1045
337 AlalysMetMetAsnGlnArgThrLysMetGluIleLysGlu..... 351
1046 .....TGCCCTGTGTGAGGGGATGCTAAAGAGAAATACTGCGAGCCACC 1089
352 .....SerSerSerp 355
1090 GGAACCTCATCCAGTCCACCCATCTGCACACTTCTCCCCCTTCAGCTT 1139
355 heAspPheArgThrSerGluIleSerAlaLysGluGluLeu..... 368
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1140 TTCCAACAGTCACTACTGTGTGGCAGGACATGATAGATACCATCCAAAG 1189
369 ...ValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGluLe 384
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1190 CCAGTGTGTCATATGGTTTTCATCAGAACACATTCAGCAGACCTCAACAG 1239
384 uAsnTyrSer..... 387
1240 AAACATATAGTAATCAACAGACTTCACGGGAAAAATGAATCAACAATTTG 1289
388 .....PheAspLysAsnAlaAspThrThrMetLysTrpGlnThrLysAla 402
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

1290 AACAGATAGATAAAAAAATTGGAACGAAATTTAAGTTTGTGAGATTAAAGA 1339
403 PheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAspLeuG1 419
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1340 GTCCCTCTCCAAGAGGGACCA.....AAAGTTTGTGAGGAA 1377
419 ySerLeuLeuPheGluGly.....C 426
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1378 CACACTTTTGAATAGGGGACATGCAATTTAAATTTAAATCTGCTTCACCTT 1427
426 ys.....SerAsnSerLysPro.....ValAsn 433
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1428 GTATAGCTGATAAAATCTTAAGCCACAGGAATTAAGTTTCAGATCTAAAT 1477
434 AlaAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSe 450
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1478 GTCGGTGATACTTCCAGAAATTTCTGTGTGGACTGCAGTGAACACAATC 1527
450 r.....ThrProPheGluLeuIleGlnArgGluThrL 462
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1528 AAACAAAGTTTCAGTTACTCCACAGAGAATCCCGAATTCAGACACAC 1577
462 ys.....GluValAspSerLysGluAsnPheSerTyr 472
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1578 CTCCAAGGCCAGACCGCTTGCTCTTGATGAGAAAGGACATGTACAGTGG 1627
473 LeuGluSerGlnProHisAspSerCysPheValGluMetGlnAlaGlnLy 489
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1628 TCATTTCATGACCTGAAAAATGCC.....ATACCCATACCTGA 1665
489 sValMethHisValSerSerAlaGluLeuAsnTyr..... 500
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1666 TTATCTGAAGGCAATTCCTCAGATATCACTATCAAACTAGGAAAACTG 1715
501 ..SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnVal 516
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1716 TGAGTTTAAACACCAAGCTCTACAACACAAGTTGAAACACCTGCTTGTG 1765
517 LysHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuVa 533
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1766 GATCATGATAACACTTCA.....CCACTCTT 1791
533 lGluAsnProTyrPheSerSerTrpProSerSerGlyThrSerSerLysM 550
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1792 CAGAACACCCCTCAGTTTACTAATCCACTCACTCTGATGACTCA.... 1837
550 etSerLeuAspLeuProGluLys...GlnAspGlyThrValPheProSer 565
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1838 .....GACTCAGATGAAAGAACTCTGATGCTGTGCTGACCCAGAAT 1879
566 SerLeuLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHi 582
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1880 AAAACTAATATTTCAACAGCAAGTGCCACAGTTTCTGCTGCCACTAGTAC 1929
582 sSerSerLeuLeuAsnSerProThrAsnIleSerSerSerLeuLeuAsnG 599
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1930 TGAAGCATTTCTACTAGGAAAGTATGCCAATGTCCATTGCTGCTAGACATA 1979
599 lNgluSerAlaValLeuAlaThrAlaProArg.....IleAsp 611
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1980 ATATAGCAGGAACAACACATCTCAGTGTCTGCTGAAAAAGATGTTGATGT 2029
612 AspGluIleProProProLeuProValArgThrProGluSerPheIleVa 628
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2030 GAAATTCACCTCTCCCTACCTACCTGAAAGAACTCCTGAACTGTTGTGTT 2079
628 lValGluGluAlaGlyGluPheSerProAsnValProLysSerLeuSers 645
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2080 AGCAAGTGAA.....CATAATACACT..... 2101
645 erAlaValLysValLysIleGlyThrSerLeuGluTrpGly..... 658
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2102 .....GTAAAGTCGGAATGGAGTGAACCTCAA 2128

```



```

1020 .....GATTCTCCTCTCCAAACACCAAGGACCCGAGT..... 1055
337 AlaLysMetMetAsnGlnGlnArgThrLysMetGluLysGlu..... 351
1056 .....TGCCTGTGTGAAGGGGATGCTAAAGAGAATACTGACGCCACC 1099
352 .....SerSerSerp 355
1100 GGAACCTCATCCAGTGCACCCATCTTGACACCTTCCTCCCTTCAGCTT 1149
355 heAspPheArgThrSerGluLysSerAlaLysGluGluLeu..... 368
1150 TTCCAACAGTCACTACTGTGTGGCAGGACAATGATAGATACCATCCAAAG 1199
369 ...ValLeuHisProAlaLysSerThrSerPheAspPheLeuGluLe 384
1200 CCAGTGTTCATATGGTTTCATCAGAACACATTCAGCAGACCTCAACAG 1249
384 uAsnTyrSer ..... 387
1250 AAACATATAGTAATCAACAGAACTTCACGGGAAAAATGAATCAACAATTG 1299
388 .....PheAspLysAsnAlaAspThrThrMetLysTrpGlnThrLysAla 402
1300 AACAGATAGATAAAAAATTGGAACGAAATTTAAGTTTGAGATTGAAGAAG 1349
403 PheProLysValGlyGluProLeuGlnLysHisGlnSerLeuAspLeuG 419
1350 GTCCCTCTCCAAAGAGGGACCA.....AAAAGTTTGTGGGAA 1387
419 ySerLeuLeuPheGluGly.....C 426
1388 CACACTTTGTAATAGGGGACATGCAATTAATAATCTCTCTCACCTT 1437
426 ys.....SerAsnSerLysPro.....Valasn 433
1438 CTATAGCTGTATAAATCTCTAAGCACAGGAATTAAGTTTCAGATCTAAAT 1487
434 AlaAlaGlyArgTyrPheAsnSerLysValProLysThrArgThrLysSe 450
1488 GTCGGTGATACTCCCAAGTTCTGTGTGGAGTCAGTGAACACAAATC 1537
450 r.....ThrProPheGluLeuGlnGlnArgGluThrL 462
1538 AAACAAAGTTTCAGTTACTCCACGAGAGATCCAGAAATTCACACACAC 1587
462 ys.....GluValAspSerLysGluAsnPheSerTyr 472
1588 CTCCAAGGCCAGACCGCTTGCTCTTGATGAGAAAGGACATGTAACGTGG 1637
473 LeuGluSerGlnProHisAspSerCysPheValGluMetGlnAlaGlnLy 489
1638 TCATTTTCATGACCTGAAATGGCC.....ATACCCATACCTGA 1675
489 sValMetHisValSerSerAlaGluLeuAsnTyr..... 500
1676 TTTATCTGAAGGCATCTCTCAGATATCACTATCAAACTAGGAAACACTG 1725
501 ..SerLeuProTyrAspSerLysHisGlnLeuArgAsnAlaSerAsnVal 516
1726 TGAGTTTAAACCAAGTCTCAACACACAAAGTTGAACACCTGATCTTGTG 1775
517 LysHisHisAspSerSerAlaLeuGlyValTyrSerTyrLysProLeuVa 533
1776 GATCATGATACACTTCA.....CCACTCTT 1801
533 lGluAsnProTyrPheSerSerTrpProSerGlyThrSerSerLysM 550
1802 CAGAACACCCCTCAGTTTACTATATCCACTTCACCTCTGATGACCTCA... 1847
550 etSerLeuAspLeuProGluLys...GlnAspGlyThrValPheProSer 565

```

```

1848 .....GACTCAGATGAAAGAACTCTGATGCTGTGTGACCCAGAAT 1889
566 SerLeuLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHi 582
1890 AAAACTAATATTCAACGCAAGTGCCACAGTTTCTGCTGCCACTAGTAC 1939
582 sSerSerSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnG 599
1940 TGAAGCATTCTACTAGGAAGATTATCCCAATGTCCATTGCTAGACATA 1989
599 lGluSerAlaValLeuAlaThrAlaProArg.....Ileasp 611
1990 ATATAGCAGGAACAACACATTCAGGTGCTGCTGAAAAAGATGTTGATGTTAGT 2039
612 AspGluIleProProLeuProValArgThrProGluSerPheIleVa 628
2040 GAAGATTACCTCTCTCCCTACCTGAAAGAACTCCTGAATCGTTTGTTGTT 2089
628 lValGluGluAlaGlyGluPheSerProAsnValProLysSerLeuSers 645
2090 AGCAAGTGAA.....CATAATACACT..... 2111
645 erAlaValLysValLysIleGlyThrSerLeuGluTrpGly..... 658
2112 .....GTAAGATCGGAATGGAGTGAACCTCAA 2138
659 .....GlyThrSerGluProLysLysPheAspAspSerValIleLeuAr 673
2139 AGTCAGGAACGATCTGAACAAAAAAGTCTCAAGGC.....TTGAT 2179
673 gProSerLysSerValLysLeuArgSerProLysSerGlyLysAsnPhe 689
2180 AACCTCTGAAATGAGAAATGTGATCATCCACGGGAGGTATTCACTAT 2228

seq_name: gb_ro:MMPTPPES

seq_documentation_block:
LOCUS MMPTPPES 2676 bp DNA linear ROD 18-SEP-1995
DEFINITION M.musculus mptp-pest gene.
ACCESSION X86781
VERSION X86781.1 GI:804997
KEYWORDS protein-tyrosine-phosphatase.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Charest,A., Wagner,J., Shen,S.H. and Tremblay,M.L.
TITLE Murine protein tyrosine phosphatase-PEST, a stable cytosolic
protein tyrosine phosphatase
JOURNAL Biochem J. 308 (Pt 2), 425-432 (1995)
MEDLINE 95289971
REFERENCE
AUTHORS Tremblay,M.L.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-1995) M.L. Tremblay, McGill University, Dept of
Biochemistry, 3655 Drummond, Montreal, Quebec H3G 1Y6, CANADA
COMMENT
FEATURES
Location/Qualifiers
source
1..2676
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/map="region A3-B"
/clone_lib="Lambda zap II"
/dev_stage="17.5 days"
misc_feature 1..14
/note="synthetic linker"
gene 330..2657
/genes="mptp-pest"
CDS 330..2657
/genes="mptp-pest"

```



```

seq_name: gb_pat:AX306312

seq_documentation_block:
LOCUS      AX306312                               2983 bp    DNA        linear    PAT 11-DEC-2001
DEFINITION Sequence 1063 from Patent WO0188188.
ACCESSION  AX306312
VERSION     AX306312.1 GI:17645552
KEYWORDS
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (sites)
AUTHORS     Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE       Method for examining ischemic conditions
            Patent: WO 0188188-A 1063 22-NOV-2001;
            School Juridical Person Nihon University (JP)
FEATURES
            Location/Qualifiers
                source          1..2983
                                /organism="Mus musculus"
                                /db_xref="taxon:10090"
BASE COUNT  923 a    634 c    681 g    745 t
ORIGIN
alignment_scores:
    Quality: 1104.50      Length: 743
    Ratio: 2.493          Gaps: 22
    Percent Similarity: 59.623    Percent Identity: 35.935
alignment_block:
US-09-600-358A-4 x AX306312
Align seg 1/1 to: AX306312 from: 1 to: 2983
1 MetAspGlnArGluIleLeuGlnLysPheLeuAspGluAlaGlnSerIly 17
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||
52 ATGGAGCAAGTGGAGATCCTCCTGAGGAGGTTCATCCAGAGGGTCCAGGCCAT 101
17 sLysIleThr.....LysGluGluPheAlaAsnGluPheLeuL 30
|||||  |||||||  |||||||  |||||||  |||||||  |||
102 GAAGAGTCGGGATCACATGGGAGGACAACTTCGCCGGGACATCATGC 151
30 ysLeuLysArgGlnSerThrLysTyrLysAlaAspLysThrTyrProThr 46
|||||  |||||||  |||||||  |||||||  |||||||  |||
152 GATTGAGAAGATTGCTACCAAAATATAGACAGAAAGATTATCCACA 201
47 ThrValAlaGluAsnAlaLysAsnIleLysLysAsnArgTyrLysAspI 63
|||||  |||||||  |||||||  |||||||  |||||||  |||
202 GCCACTGGAGAAAAGAAAATGTTAAAAGAACACATATAGGACAT 251
63 eLeuProTyrAspTyrSerArgValGluLeuSerLeuIleThrSerAspG 80
|||||  |||||||  |||||||  |||||||  |||||||  |||
252 ACTGCCATTGATCACACCCGACATTAAGTTGACTTTGAAGACTCCATCCC 301
80 luAspSerSerTyrIleAsnAlaAsnPheIleLysGlyValTyrGlyPro 96
|||||  |||||||  |||||||  |||||||  |||||||  |||
302 AAGATTTCAGATTATATCAATGCAAAATTTATTAAAGGTGTGTATGGGCA 351
97 LysAlaTyrIleAlaThrGlnGlyProLeuSerThrThrLeuLeuAspPh 113
|||||  |||||||  |||||||  |||||||  |||||||  |||
352 AAAGCATATGTGGCAACCAAGGCGCTTTCGGGAATACAGATCATAGACTT 401
113 eTrpArgMetIleTrpGluTyrSerValLeuIleValMetAlaCysM 130
|||||  |||||||  |||||||  |||||||  |||||||  |||
402 CTGGAGGATGATATGGGAGTATATATGTTGATGATCGTGATGGCCTGTC 451
130 etGluTyrGluMetGlyLysLysLysCysGluArgTyrTrpAlaGluPro 146
|||||  |||||||  |||||||  |||||||  |||||||  |||
452 GAGATTGAGATGGGAGGAAAAGTGTGAGCGCTACTGGCCCTTGTGAT 501
147 GlyGluMetGlnLeuGluPheGlyProPheSerValSerCysGluAlaG 163
|||||  |||||||  |||||||  |||||||  |||||||  |||

```

502 GGAGAGATCCTATAACATTTGGACCAATTTAAATTTCTTGTGAAATGA 551
163 uLysArgLysSerAspTyrIleIleArgThrLeuLysValLysPheAsnS 180
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
552 ACAAGCAAGAACGACTACTTCATCCGACACACTTTACTTGAATTCAAA 601
180 erGluThrArgThrIleTyrGlnPheHisTyrLysAsnTrpProAspHis 196
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
602 ATGAATCCCGCTCGCTCTATCAGTTTCATACGTGAACCTGGCCAGACCAT 651
197 AspValProSerSerIleAspProIleLeuLysValLysValLysVal 213
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
652 GATGTTCTTCTTCGTCATTTGATTTCTTCTGGACATGATAAGCTTAATG 701
213 qCysTyrGlnGluAspSerValProIleCysIleHisCysSerAlaG 230
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
702 GAATACCAAGACATGAGATGTGCTTATTTGATTCATTCAGTGCCTG 751
230 lyCysGlyArgThrGlyValIleCysAlaIleValAspTyrThrTrpMet 246
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
752 GCTGTGGACGAACAGGTGCTATTGTGTGCCATA..GATTACACGTGGAAC 798
247 LeuLeuLysAspGlyIleIleProGluAsnPheSerValPheSerLeuI 263
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
799 TTACTGAAGACAGGGAATTTCCAGAGGAATTTAATGTATTTAATTAAT 848
263 eArgGluMetArgThrGlnArgProSerLeuValGlnThrGlnGluGln 280
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
849 ACAAGAAATGAGAACACAGAGGCACTCGGCAGTACAAACAAGGAGCAGT 898
280 yrGluLeuValTyrAsnAlaValLeuGluLeuPheLysArgGlnMetAsp 296
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
899 ATGAACCTGTTCATAGGCTATTGCTCAACTGTTGAAACACAGCTACAAC 948
297 ValIle..... 298
949 TGTATGAATTCATGGAGCACAGAGATCCGTCGATGTAATGAATTAACC 998
299 303
999 ACTGGAATATGTCAGTTCCATTCATGATAGCGAGACGAGACCTCTCCGCCA 1048
304 305
1049 AAGCCGCGCGGACTCGAAGTTCCTTGTAGAAGGGATGCCAAGGAAGA 1098
306 GluSerGlnAlaLysHisCysIleProGluLysAsnHisThrLeuGlnAl 322
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1099 AATACTACAGCCACCAAGACCTCACCCGGTGCCACCATCTGACGCCAT 1148
322 aAspSerTyrSerProAsnLeuProLysSerThrThrLysAlaAlaLysM 339
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1149 CACCTCCTTCAGC.....CTTCCACAGTACCACCTGCTGTGAGGACAG 1192
339 et.....MetAsn 341
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1193 TGACAGGTACACCCCAAGCCAGTGCATGCACATGGCCCTCACCAGAGCAAC 1242
342 GlnGlnArgThrLysMetGluIleLysGluSerSerPheAspPheAr 358
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1243 ACCCAGCCGACCTCAACAGAA.....GCTATGATAAATCAGCG 1280
358 qThrSerGluIleSerAlaLysGluGluLeuValLeuHisProAlaLysS 375
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1281 GACC.....AATGGGGCAAAA 1297
375 erSerThrSerPheAspPheLeuGluLeuAsnTyrSerPheAspLysAsn 391
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1298 GTGAATCAGCTATTGACACATA.....GATAAGAAG 1329
392 AlaAspThrThrMetLysTrpGlnThrLysAlaPheProIle..... 405
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1330 TTAGAGCGCAATTAAGTTTGTAGATTAAAGAAAGTCCTCTCCAAGAAGG 1379

405 405
1380 GCCCAAAAGTTTTGATGGGAACACACTCTTGAANTAGGGACATGCGGATTA 1429
406ValGlyGluProLeuGlnLysHisGln 414
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1430 AAATTAATTCGTCTTCATCTCTCTAGTTGACAGAAACCTCTAAACCACAG 1479
415 SerLeuAspLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysPr 431
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1480 GAGTTAAGTGCAGGTGCCCTAAAGGTGATGATGTATCTCAGAAATCTTG 1529
431 oValAsnAlaAlaGlyArgTyr..... 438
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1530 CGCGGACTGTAGTGGCGCTCATTCACACAGAGCTGTGAGTCGCAGAGG 1579
439PheAsnSerLysValProIleThrArgThrLysSerThrPro 452
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1580 AGTCCCGACGAGCAACTCACACACACT...CCAGGGCCAGACTGCTGCCT 1626
453 PheGluLeuIleGlnArgGluThrLysGluValAspSerLysGluAs 469
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1627 CTCGATAAG...AAAGACACGTAAACGTGCTCATTATGGACCTGAAAA 1673
469 nPheSerTyrLeuGluSerGlnProHisAspSerCysPheValGluMetG 486
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1674 TGCCACACCTGTACCCGACTCACT...GACGGCAAAATCCCAGATAATC 1720
486 lnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyrSerLeu 502
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1721 ATTCTCAGACTCTGAAACCGTGAGTTCACACCC.....AACTCCACC 1764
503 ProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnValLysHisH 519
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1765 GCAGAGAGGAGGCCACGATCT.....ACAGAGCACCA 1799
519 sAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValGluAsnP 536
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1800 CAACAGCTCC.....CCTCTGTGAAAGCTC 1825
536 ro.....TyrPheSerSerTrpProSerGly 545
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1826 CCTCAGCTTTACCAACCCCTCTTCACTCTGACGACTGGCAGCTCAGAGGA 1875
546 ThrSerSerLysMetSerLeuAspLeuProGluLysGlnAspGlyThrVa 562
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1876 GGGAGCTCT.....GATGTCCTGT 1895
562 lPheProSerSerLeuLeuProThrSerSerThrSerLeuPheSerTyrT 579
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1896 GACCAGGAACAAAACATAGCATTTCAACAGCAAGTGGCCACAGTGTCTCCTG 1945
579 yrAsnSerHisSerSerLeuSerLeuAsnSerProThrAsnIleSerSer 595
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1946 CCAGTAGTGTGAGAGT...GCTTGCCATAGGAGAGTATTCGCCGATGCC 1992
596 LeuLeuAsnGlnGluSerAlaValLeuAlaThrAlaProArg..... 609
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
1993 ATTGCCAGACAGGAA.....GTAGAGGACGCCGCGCATTCAGGTGC 2033
610IleAspAspGluIleProProLeuProVala 621
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
2034 TGAGAAAGATGCTCATGTTAGTGAGGAGTCCCTCTCTCTTACCTCAAC 2083
621 rgThrProGluSerPheIleValValGlu 630
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
2084 GAACCTCTGAGTCTCTTGTATTAGCAGAT 2112

seq_name: gb_ro:MMP19PTP

seq_documentation_block:

LOCUS MMP19PTP

2983 bp

mRNA

linear

ROD 15-NOV-1993


```

1049 AAGCCGCGGGGACTCGAAGTTGCTTGTAGAAAGGGGATGCCAAGGAAGA 1098
|||||
306 GluSerGlnAlaLysHisCysIleProGluLysAsnHisThrLeuGlnAl 322
|||||
1099 AATACTACAGCCACCAACCACTCACCGGTGCCACCCCATCTCAGCCCAT 1148
|||||
322 aAspSerTyrSerProAsnLeuProLysSerThrThrLysAlaAlaLysM 339
|||||
1149 CACCTCCTTCAGC.....CTTCCACAGTTTACCACCTGTGTGAGGACAG 1192
|||||
339 et.....MetAsn 341
|||
1193 TGACAGGTACCAACCAAGCCAGTGTGCACATGGCCTCACAGAGCAAC 1242
|||||
342 GlnGlnArgThrLysMetGluLysGluSerSerSerPheAspPheAr 358
|||||
1243 ACCAGCGGACCTCAACAGAA.....GCTATGATAAATCAGCG 1280
|||||
358 gThrSerGluIleSerAlaLysGluGluLeuValLeuHisProAlaLys 375
|||||
1281 GACC.....AATGGGCAAAA 1297
|||||
375 erSerThrSerPheAspPheLeuGluLeuAsnTyrSerPheAspLysAsn 391
|||||
1298 GTGAATCAGCTATTGACACATA.....GATAAGAAG 1329
|||||
392 AlaAspThrThrMetLysTyrGlnThrLysAlaPheProIle..... 405
|||||
1330 TTAGAGCGCAATTTAAGTTTGTAGATTAGAAGTCCCTCTCCAAGAAGG 1379
|||||
405 ..... 405
1380 GCCCAAAAGTTTGTAGGAACACACTCTTGAATAGGGGACATCGGATTA 1429
|||||
1430 AAATTAAATCTGCTTCATCTCTGTAGTTGACAGAACCTCTAAACCCACAG 1479
|||||
415 SerLeuAspLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysPr 431
|||||
1480 GAGTTAAGTCAGGTGCCCTTAAAGGTTTGATGATGATATCTCAGAATCTTG 1529
|||||
431 oValAsnAlaAlaGlyArgTyr..... 438
|||||
1530 CGCGGACTGTAGTCGCGCTCATTCACACAGAGCTGTGAGTCGTCACAGG 1579
|||||
439 .....PheAsnSerLysValProIleThrArgThrLysSerThrPro 452
|||||
1580 AGTCCCAAGCACTACACACACCT...CCACGGCCAGACTGTGTCCT 1626
|||||
453 PheGluLeuIleGlnArgGluThrLysGluValAspSerLysGluAs 469
|||||
1627 CTCGATTAAG...AAAGACAGTAACTGCTGTCATTCATGACCTGAAGA 1673
|||||
469 nPheSerTyrLeuGluSerGlnProHisAspSerCysPheValGluMetG 486
|||||
1674 TGCCACACCTTACCCGACCTACCT...GACGGCAATCCACAGATAATC 1720
|||||
486 lnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyrSerLeu 502
|||||
1721 ATTCTCAGACTCTGAAACACGCTGAGTTCACACCC.....AACTCCACC 1764
|||||
503 ProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnValLysHisH 519
|||||
1765 GCAGAAGAGGAGCCCGAGTCTT.....ACAGAGCACCA 1799
|||||
519 aspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValGluAsnPr 536
|||||
1800 CAACAGCTCC.....CCTCTGTGAAGCTC 1825
|||||
536 ro.....TyrPheSerSerThrProProSerGly 545
|||||

```

```

1826 CCCTCAGCTTTACCAACCCCTTCTTCACTCTGACGACTGGCACTCAGACGGA 1875
|||||
546 ThrSerSerLysMetSerLeuAspLeuProGluLysGlnAspGlyThrVa 562
|||||
1876 GGAGGCTCT.....GATGGTGTCTGT 1895
|||||
562 lPheProSerSerLeuLeuProThrSerSerThrSerLeuPheSerTyrT 579
|||||
1896 GACCAGGAACAAAACAGCATTTTCAACAGCAAGTGCACAGCTGTCTCCTG 1945
|||||
579 yrAsnSerHisSerSerLeuSerLeuAsnSerProThrAsnIleSerSer 595
|||||
1946 CCAGTAGTGTGAGAGT...GCTTGCCATAGGAGAGTATTGCCGAGTCTCC 1992
|||||
596 LeuLeuAsnGlnGluSerAlaValLeuAlaThrAlaProArg..... 609
|||||
1993 ATTGCCAGACAGGAA.....GTAGCAGGCACGCCGATTCAGGTGC 2033
|||||
610 .....IleAspAspGluIleProProLeuLeuProVala 621
|||||
2034 TGAGAAAGATGCTGATGTTAGTGAGGAGTGCCTCTCTTACCTGAAC 2083
|||||
621 rgThrProGluSerPheIleValValGlu 630
|||||
2084 GAATCCTGAGTCTTTTGTATTAGCAGAT 2112
|||||
seq_name: gb_ro:RATRKPTP
seq_documentation_block:
LOCUS RATRKPTP 1608 bp mRNA linear ROD 08-FEB-1999
DEFINITION Rat mRNA for protein tyrosine phosphatase, complete cds.
ACCESSION D38072
VERSION D38072.1 GI:567262
KEYWORDS protein tyrosine phosphatase.
SOURCE Rattus norvegicus (strain:Sprague-Dawley) kidney cDNA to mRNA,
clone:RKPTP.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;
Rattus
REFERENCE 1 (bases 1 to 1608)
AUTHORS Moriyama,T.
TITLE Direct Submission
JOURNAL Submitted (26-AUG-1994) Toshiki Moriyama, Osaka University Medical
School, 1st Dept. of Medicine; 2-2 Yamadaoka, Suita, Osaka 565,
Japan (Tel:06-879-3632, Fax:06-879-3639)
REFERENCE 2 (bases 1 to 1608)
AUTHORS Moriyama,T., Kawanishi,S., Inoue,T., Imai,E., Kaneko,T., Xia,C.,
Takenaka,M., Noguchi,T., Kamada,T. and Ueda,N.
TITLE cDNA cloning of a cytosolic protein tyrosine phosphatase (RKPTP)
from rat kidney
JOURNAL FEBS Lett. 353 (3), 305-308 (1994)
MEDLINE 95046282
FEATURES
source
1..1608
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/gb_xref="taxon:10116"
/clone="RKPTP"
/tissue_type="kidney"
135..1283
/codon_start=1
/product="protein tyrosine phosphatase"
/protein_id="BAA07266.1"
/db_xref="GI:699627"
/translation="MEQVEILRRFTQVQAMKSPDHNGEDNPFARDFMLRLRLSTKYRT
EKIYPTATGEKENYKKNRYKDILFDHSRVKLTLLKTPSQSDIYNANFIKGVYGPRA
VYATQPLANTVIDFWIMTEYNVYIIVMACREFEMGRKKKERYWPLXGDDITPAPE
KISCENEQATDYFTIRLLLEFQNSRRLYOFHYVWPDHVPSSDFSLDMLSMRK
YOEHEVDPTICHSAGCGRTGAICAIDYTNLLKAGKIPPEENVNLQEMRTORHSA
VOTKEQYELVHRAIQLEKQLOLVEIHGAOKITDGNISTGNMYSSIDSEKQDPPPP
KPPTRTSCLVGDGAKEEILQPPPEHPVPIPLPSPSPSPAPPLHLCVARQ"
BASE COUNT 490 a 347 c 410 g 361 t

```

ORIGIN

alignment_scores:
 Quality: 1058.00 Length: 465
 Ratio: 3.112 Gaps: 12
 Percent Similarity: 73.118 Percent Identity: 47.742

alignment_block:
 US-09-600-358a-4 x RATRKPTP ..

Align seg 1/1 to: RATRKPTP from: 1 to: 1608

```

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerly 17
135 ATGGAGCAAGTGGAGATTCTGAGGAGGTTCAATCCAGAGGGTCCAGGCCAT 184
17 slvsylethr.....LysGluGluPheAlaAsnGluPheLeuL 30
185 GAAAGTCCGGACCACAATGGGGAGGACAACATTCGCCGGGACTTCATGC 234
30 ysLeuLysArgGlnSerThrLysTyrLysAlaAspLysThrTyrProThr 46
235 GATTGAGAAGATTGTCTACCAATATAGAACAGAAAAAGATTATCCACA 284
47 ThrValAlaGluAsnAlaLysAsnIleLysLysAsnArgTyrLysAspI 63
285 GCCACTGGAGAAAGAGAAATGTTAAAGAACAGATATAAGGACAT 334
63 eLeuProTyrAspTyrSerArgValGluLeuSerLeuIleThrSerAspG 80
335 ACTGCCATTGTATCAGCAGCGAGTTAAGTTGACTTTGAAGACACCATCC 384
80 luAspSerSerTyrIleAsnAlaAsnPheIleLysGlyValTyrGlyPro 96
385 AAGATTACAGATTATATCAATGCMAATTTATTAAGGGTGTATATGGCGC 434
97 LysAlaTyrIleAlaThrGlnGlyProLeuSerThrThrLeuLeuAspPh 113
435 AGAGCATACGTGGCAACCCAGGGCCTTTGGGAATACAGTCATAGACTT 484
113 eTrpArgMetIleThrGluTyrSerValLeuIleValMetAlaCysM 130
485 CTGGAGGATGATATGGGAGTACAATGTGGTGATCATCTTAATGGCCTGC 534
130 eTGlutyrGluMetGlyLysLysCysGluArgTyrTrpAlaGluPro 146
535 CTGAATTTCAGATGGGAAGGAAAAAGTGTGAGCGCTATTGGCCTTTGTAT 584
147 GlyGluMetGlnLeuGluPheGlyProPheSerValSerCysGluAlaG 163
585 GGAGAAGATCTCTAATCAATTTGCACCATTTAAATTTCTTGTGAAAAATGA 634
163 uLysArgLysSerAspTyrIleIleArgThrLeuLysValLysPheAsnS 180
635 ACAACGAAGACAGACTACTTCATTCGACACTTTTACTTGNATTTCAA 684
180 eRGlutThrArgThrIleTyrGlnPheHisTyrLysAsnTrpProAspHis 196
685 ATGAATCCCGTCTGACTCTATCAGTTTCACTACGTGAAGTGGCCAGACCAT 734
197 AspValProSerIleAspProIleLeuGluLeuIleTrpAspValar 213
735 GATGTCCCTCTGCTCATCTCTATCTGACATGATAAGTCTAATGAG 784
213 gCysTyrGlnGluAspAspSerValProIleCysIleHisCysSerAlaG 230
785 GAAGTACCAAGACAGAGAGATGTGCCTATTGTATCCATTCAGTGCAG 834
230 lYcysGlyArgThrGlyValIleCysAlaIleValAspTyrThrTrpMet 246
835 GCTGTGGAGGACAGGTGCTATTTGTGCCATC...GACTACACGTGGAAC 881

```

```

247 LeuLeuLysAspGlyIleIleProGluAsnPheSerValPheSerLeuI 263
882 TTACTGAAAGCAGGAAATTCAGAGGAATTAATGATATTAAATTAAT 931
263 eArgGluMetArgThrGlnArgProSerLeuValGlnThrGlnGluGln 280
932 ACAAGAAATGAGAACACAAAGCACTCTGCTAGTACAAACGAAGGAGCAGT 981
280 yrGluLeuValTyrAsnAlaValLeuGluLeuPheLysArgGlnMetAsp 296
982 ATGAACCTGCTCCATAGAGCTATTGCCCACTGTTTGAAGAAACAGCTACA 1031
297 Val.....IleArgAspLysHis..... 302
1032 CTGTATGAATCCACGGAGCCAGAAATCACTGATGTAATGAATTAAGTAG 1081
303 .....SerGlyThrGluSerGlnAlaLysHisCysIleP 314
1082 CACTGGAAACATGGTCAGTTCATTTAGTAGTGAACAAACAAAGATTCTCTC 1131
314 roGluLys. AsnHisThrLeuGlnAlaAspSerTyrSerProAsnLeuPr 330
1132 CTCCAAAGCCACCGGAGCTCGAAGTT...GTCTTTGTAGAGGGGATGCC 1178
330 oLysSerThrThrLysAlaAlaLysMetMetAsnGlnGlnArgThr.... 345
1179 AAGGAAGAAATCCTCCAGCCACGAGGCTCACCGGTGCCGCCCATCCT 1228
346 .....LysMetGluIleLysGluSerSerPheAspPheArgThrSer 360
1229 GACACCGTCGCTCTCTTCAGCATTCCTCAACAGTTACACTGTGTGGCAAGAC 1278
361 GluIleSerAlaLysGluGluLeuValLeuHis.....ProAlaLys 374
1279 AGTGACAGGTACCAACCAAGCCAGTGTGCACATGGCTTCCCCCAGAGCA 1328
374 sSerSerThrSerPheAspPheLeuGluLeuAsnTyrSerPheAspLysA 391
1329 GCCCCCAACG.....GACCTCAACAGAACTATGATAAGT 1363
391 snAlaAsp.....ThrThrMet 396
1364 CAGCGGACCTAATGGGGAGAACGGAATCTGCTGTTGAGCACACAGATAAA 1413
397 LysTrp...GlnThrLysAlaPheProIleValGlyGluProLeuGlnLy 412
1414 AAGTTAGAACAAAAATTTAAGTTTGAAGATTAAAGAAAGTCCCTCTCCAAGA 1463
412 s...HisGlnSerLeuAspLeuGlySerLeuLeuPheGluGly 425
1464 AGGGCCCAAAAGTTTTGTGGAACACACACTCTTGAATAGGGGA 1506

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2002, 07:00:23 ; Search time 95.32 seconds
(without alignments)
6071.260 Million cell updates/sec

Title: US-09-600-358A-3
Perfect score: 2356
Sequence: 1 tccctcaacctacttataga.....cctagtcgagacatccgg 2356

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2074.6	88.1	3580	4	US-09-081-345-1
2	216.4	9.2	1529	4	US-08-821-278A-1
3	214.2	9.1	2226	3	US-08-951-260A-1
4	130.6	5.5	2309	1	US-08-036-210-10
5	130.6	5.5	2309	2	US-08-449-609-10
6	130.6	5.5	2692	1	US-08-036-210-14
7	130.6	5.5	2692	2	US-08-449-609-14
8	130.6	5.5	3973	1	US-08-036-210-21
9	130.6	5.5	3973	2	US-08-449-609-21
10	120	5.1	466	4	US-08-821-278A-16
11	116.8	5.0	5117	3	PCT-US95-05512-1
12	116.8	5.0	5117	5	US-08-854-585-1
13	115.4	4.9	3969	2	US-08-241-853-1
14	115.4	4.9	3969	2	US-08-850-917-1
15	106.2	4.5	247	1	US-08-594-031-74
16	105.6	4.5	2173	1	US-08-036-210-6
17	105.6	4.5	2173	2	US-08-449-609-6
18	104.8	4.4	4752	1	US-08-201-697-3
19	104.8	4.4	4799	1	US-08-201-697-6
20	104.8	4.4	4815	1	US-08-201-697-5
21	100.6	4.3	5455	1	US-08-342-930-1
22	98.4	4.2	4651	1	US-08-449-644-4
23	98.4	4.2	4651	2	US-08-087-244A-4
24	90.4	3.8	5690	2	US-08-447-464-2
25	90.4	3.8	5690	2	US-08-716-679-2
26	90	3.8	6000	1	US-08-348-006B-6
27	90	3.8	6000	2	US-08-800-825A-6

28	90	3.8	6000	4	US-09-158-657-6	Sequence 6, Appli
29	90	3.8	6000	5	PCT-US94-10166-6	Sequence 6, Appli
30	88.8	3.8	5679	1	US-08-201-697-1	Sequence 1, Appli
31	87.2	3.7	5769	1	US-08-652-971-1	Sequence 1, Appli
32	87.2	3.7	5769	2	US-08-991-258A-1	Sequence 1, Appli
33	87.2	3.7	5769	2	US-08-769-399-1	Sequence 1, Appli
34	87.2	3.7	5769	3	US-08-991-253A-1	Sequence 1, Appli
35	84	3.6	2351	2	US-08-548-159-6	Sequence 6, Appli
36	80.8	3.4	6924	1	US-08-015-973-2	Sequence 1, Appli
37	80.8	3.4	6924	2	US-08-448-164-2	Sequence 2, Appli
38	78.8	3.3	4338	1	US-08-015-986A-1	Sequence 1, Appli
39	78.8	3.3	4338	2	US-08-446-363-1	Sequence 1, Appli
40	74.6	3.2	963	3	US-08-964-308-1	Sequence 1, Appli
41	74.6	3.2	963	3	US-08-964-308-7	Sequence 7, Appli
42	74.6	3.2	963	3	US-08-964-313-1	Sequence 1, Appli
43	74.6	3.2	963	3	US-08-964-313-7	Sequence 7, Appli
44	74.6	3.2	963	4	US-09-069-138-1	Sequence 1, Appli
45	74.6	3.2	963	4	US-09-069-138-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-081-345-1
; Sequence 1, Application US/09081345
; Patent No. 6228641
; GENERAL INFORMATION:
; APPLICANT: Bahija Jallal
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: PTP04 RELATED DISORDERS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/081,345
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/047,222
; FILING DATE: May 20, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 234/253
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEX: (213) 955-0440
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3580 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-081-345-1

Query Match 88.1%; Score 2074.6; DB 4; Length 3580;
Best Local Similarity 99.7%; Pred. No. 0;

	Matches	2090;	Conservative	0;	Mismatches	4;	Indels	3;	Gaps	1;			
QY	1	tcctcaac	tactattagac	tattttctgtc	gtcgcagcatg	accacaaagaa	attc	60					
Db	12	TCCCTCA	ACCTACTTATAG	ACTATTTTCTTGCTCT	GCAGCATGG	ACCAAGAAATTC	71						
QY	61	tgagaa	gttctctggatg	agggcccaagc	agaagaaattact	aaagagaggtt	gccaatg	120					
Db	72	TGCAGAA	GTTCCTGGATG	AGGCCCAAAGCA	AAATTTACTTAA	AGAGAGGTTT	GCCAAATG	131					
QY	121	aattctg	agsgtgaaag	gcaatctcc	aaagtcac	aggcgagacaa	acctctac	caa	180				
Db	132	AATTTCT	GAAAGCTTCT	CCAAAGCAATCT	CCAAAGTCA	AGGCGAGACAA	AAACCTTAT	CCTACAA	191				
QY	181	ctg	ggctgaga	tgc	caagaata	caagaaaa	caaatata	laagatattt	gccctatg	240			
Db	192	CTGTGG	CTGAGAGCC	CAAGAAATCA	AGAAACAGAT	TAAGGATATTT	TGGCCTATG	251					
QY	241	attatag	ccgggtaga	actatccct	gataac	ctctgatg	agattcc	cagctacat	caatg	300			
Db	252	ATTATAG	CCGGTAGAA	CTATCCCTG	ATAACTCTG	ATGAGGATTC	CCAGCTAC	ATCAATG	311				
QY	301	ccaa	ctcattaa	ggggagtt	tatgga	cccaagctt	atatatgc	ccaccagg	gtcccttat	360			
Db	312	CCAACTT	CTAATAAGG	AGTTTATGG	ACCAAGGCTT	ATATTTGCC	ACCCAGGGT	CCTTTAT	371				
QY	361	ctaca	ccctcctg	actctcg	agagatg	attggg	aatatag	tgtccttat	catgtta	420			
Db	372	CTACA	ACCCCTCTG	CACTTCTGG	AGGATGATTT	GGGAATATAG	TGTCTTAT	CATTGTTA	431				
QY	421	tgg	catgcatg	gagatg	aaatgg	aaagaaa	agtgtgag	cgtactg	ggctgag	ccag	480		
Db	432	TGGCAT	CGATGGAGT	ATGAATG	GAAAGAAAGT	GTCGACGCT	TACTGGCT	GAGCCAG	491				
QY	481	gagaga	tgcagctg	gaatttgg	cccttctc	tgtatcc	ctgtga	agctg	aaagaaat	540			
Db	492	GAGAGAT	GCAGCTGGA	ATTTGG	CCCTTCTCTGT	ATCTGTG	TAAAGTGA	AAAGTGA	AAAGAAAT	551			
QY	541	ctg	attata	taacag	actctaa	agtta	agttca	atag	tgaactc	gaactat	600		
Db	552	CTGATTA	TATANTCAG	AGCTCTAA	AGTTAAGTTCA	ATAGTGA	AAAGTGA	AAAGTGA	AAAGTGA	611			
QY	601	agtt	tcattaca	gaatttgg	ccagaccat	gatgtac	cttcat	ctatag	accctat	ctctg	660		
Db	612	AGTTTCA	TATACAAG	ATTTGGC	CAGACCATG	ATGACCTT	CACTATAG	ACCCATTTCT	CTTG	671			
QY	661	agctca	tctggga	tgtac	gttgtacc	aaaggga	tgcag	tgttcc	catatgc	atcact	720		
Db	672	AGCTTCA	CTGGGATG	TACGTTT	TACCAAGAG	GATGAC	AGTGTTCC	CCATATG	CATTTC	ACT	731		
QY	721	gcagt	ctgctgtg	gaagcag	ctggttatt	tgtgctatt	gttgatt	gattacat	gatgatg	atg	780		
Db	732	GCA	GCTGCTG	TGGAGG	AGTGGTGTT	TGTGCTA	---	TGATATAC	TGGATGT	788			
QY	781	tgtc	taaaaga	tgggata	taattc	ctgaga	acttcag	tgttttc	agttg	atcccg	ggaatgc	840	
Db	789	TGCTAAA	AGATGGCA	TAAATCCT	CAGAACTT	CAGTGT	TTTTTTCAG	TTTGAT	CCGGGAAATGC	848			
QY	841	ggaca	cgagggc	cttcattag	tccaac	gcgagg	aaacaata	tga	aactg	gtctac	aatgctg	900	
Db	849	GGACAC	AGAGCCCTT	CATTAGT	TTCAAAGC	AGAAACAA	TATGA	CTGCTT	ACAAATGCTG	908			
QY	901	tatt	agaactatt	taagag	acagatg	gatgtt	atcag	agata	aaacatt	ctc	ggaacag	960	
Db	909	TATTAG	AACATATTTA	AGAGACAG	ATGGATGTT	TATCAG	AGATTA	AAACATTTCT	TGGA	CAGAGA	968		
QY	961	gtca	gcaaa	gcatg	tattctg	agaaaaat	cacact	ctcca	gcag	actctt	attctc	1020	
Db	969	GTC	AGCA	AAAGCAT	TGTTATTC	TGAAAAAT	CACACTCT	CCA	AGCAG	ACTCTT	ATTCTC	1028	
QY	1021	cta	attacc	aaag	tacc	aaag	cagca	aaatgat	ga	acca	caaaag	gacaaa	1080
Db	1029	CTAATTT	TACAAAAAGT	TACCAAA	AGAGAGCA	AAAAATG	ATGAAC	CAACAA	AGGAC	AAAAA	1088		

RESULT 2
US-08-821-278A-1
; Sequence 1, Application US/08821278A

; Patent No. 6238902
; GENERAL INFORMATION:
; APPLICANT: Cheng, Jill
; APPLICANT: Lasky, Laurence A.
; TITLE OF INVENTION: Protein Tyrosine Phosphatases
; FILE REFERENCE: P1010R1
; CURRENT APPLICATION NUMBER: US/08/821.278A
; CURRENT FILING DATE: 1997-03-20
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 1
; LENGTH: 1529
; TYPE: DNA
; ORGANISM: Mus musculus
US-08-821-278A-1

Query Match 9.2%; Score 216.4; DB 4; Length 1529;
Best Local Similarity 57.4%; Pred. No. 1.9e-50;
Matches 410; Conservative 0; Mismatches 301; Indels 3; Gaps 1;
Qy 200 gaatatcaagaaacagataaggaattttgcccctatgattatagccgggtagaact 259
Db 183 gaacacgaacgaacgcgcacaaagatggtgagcatatgatgagacaagagcctcct 242
Qy 260 atccctgataacccctctgatgaggattccagctacatcaatgccaaacttcattaaaggagt 319
Db 243 ttccctgctccaagagagagagacatgaattacatcaatgccaaacttcacgggcat 302
Qy 320 ttatggaccacagcttatattgcccacccagggtccctttatctacaaccctccctggactt 379
Db 303 agatggagccagcgcctacattgcagcgaagagaccctgctcacacactgttggactt 362
Qy 380 ctggagatgattggaaatagttcccttatctattgttgcagtcgcatgagatga 439
Db 363 ctggcgcctggttggaggttggggccaagggttaactctgaggtccctgtccaagagacaga 422
Qy 440 aatgggaagaaagtgtagcgcctactggtgagccagagagatgcagctgggaatt 499
Db 423 aaatggagcgaagagtgtaacgctattggccgggagcagagcctctaaagctgg 482
Qy 500 tggccctctctctgtatccctgtgaagctgaaaaagaaatctgtattatataatcaggac 559
Db 483 gcctttctgcatccctgcacaaaggagacaaactgaatgcagacatcactctcaggac 542
Qy 560 tctaaaagttcaagttaagtgaactgaactatcactatcactgatttattacaagaattg 619
Db 543 cctccaggtttacattccagaaggaattccgctctgtgcaccaactacagtatatgtcctg 602
Qy 620 gccagaccatgatgtacattcattatagaccctattcttgagctcactcgtgggtatcag 679
Db 603 gccagaccaggggttcccagcaggttctgtatcactatcaccattcgttggaggagcccg 662
Qy 680 ttgttaccagagagatgacaggtttcccattatgcattcattcagtgctgctgctggaag 739
Db 663 ctgctccaaaggcttggaacctgaccctctgtgtccactcagtgctgctgctgagc 722
Qy 740 gactggtgtattgtgctattgttattatatacatgatgttgcataaagatgggataat 799
Db 723 aacaggttctctgtgcg---tggtgactatgtgagcagtgctgctgacccagacaat 779
Qy 800 tcttgagaactcaggtttctgattgtatccgggaatgcgacacagagccttcatt 859
Db 780 ccttcccaactcagctctctccaagtggtcctggagatgcggaacagcggcctgcagc 839
Qy 860 agttcaacacgaggaacaaatgaactggctcacaatgctattagaaactatt 913
Db 840 agtcagacagagagcagctacaggttccctgtaccacacacagtcgctcagctatt 893

RESULT 3
US-08-951-260A-1
; Sequence 1, Application US/08951260A
; Patent No. 6004791

; GENERAL INFORMATION:
; APPLICANT: Aoki, Naohito
; APPLICANT: Ulrich, Axel
; TITLE OF INVENTION: PROTEIN TYROSINE PHOSPHATASE PTP20
; TITLE OF INVENTION: AND RELATED PRODUCTS AND METHODS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/951,260A
; FILING DATE: October 16, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,860
; FILING DATE: No 6004791ember 13, 1996
; APPLICATION NUMBER: PCT/1897/00946
; FILING DATE: June 17, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 227/004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2226 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA to lRNA
US-08-951-260A-1

Query Match 9.1%; Score 214.2; DB 3; Length 2226;
Best Local Similarity 57.4%; Pred. No. 9.3e-50;
Matches 406; Conservative 0; Mismatches 298; Indels 3; Gaps 1;
Qy 207 aagaaacagatataaggatattttgccttatgattatagccgggtagaactatccctg 266
Db 199 AAGAAGAACCGCTACAAAGACGTGTACCGGTATGATGAGACGAGATCATCTTTTCCTG 258
Qy 267 ataacctgatgaggattccagctacatcaatgccaacttcattaaggagattatgga 326
Db 259 CTCAGGAGGAAGGACACGGAGATTACATTAATGCCAACTTCACCGGGGCACAGATGGA 318
Qy 327 cccaagcttatattgccaccagggtccctttattctacaaccctcctggactctggagg 386
Db 319 AGCCAGCCCTACATTGGCAGCAAGGACCCCTGCCTCACACTCTGTGGACTCTTGGCGC 378
Qy 387 atgattgggaataatagttcctcttatctattgttgcagtcagtcagtgatgaatggga 446
Db 379 CTGTTTGGGAGTTTGGAAATCAAGTGTATCTGTATGGCTGTACAGGAGACAGAAATGGA 438
Qy 447 aagaaaagtgtagcgtactactggcagccagagagatgagtcagtggaatttggcct 506
Db 439 CGGAGGAAGTGTGAACCTACTTGGGCCACAGGAGCGGGAGCTCTACAGGCGCGGCTTTC 498
Qy 507 ttctctgtatcctgtgaagctgaaaaaggaactctgattataatcaggactctaaaa 566

Db 499 TGCATCACCTGACAAAGGAGACAGCACTGACTTCGGACATCACTCTCAGGACCCCTCCAG 558
QY 567 gtaagttcaatagtagaaactcgaactatctaccagtttcattacaagaattggccagac 626
Db 559 GTTACATTCAGAAGAAATCCCGTCTGTGACACAGCTACAGTACATGCTTTGGCCGGAC 618
QY 627 catgatgaccttcattctatagaccctattcttgagctcatctgggatgacgtatgctgttac 686
Db 619 CACGGGTTCCACAGAGTTCCGATCACATCTCACCATGGTGAGGAGGCCGCTGCCCTC 678
QY 687 caagaggatgacagtggtcccatatgcatctcactcagctcagtgctggtgtggaaggactgt 746
Db 679 CAAGGACTTGGACCTGGACCCCTCTGTGTCCACTGCACTGCTGGCTGTGGAGCAACAGGT 738
QY 747 gttattgtgctatggtatatacatggtatgctgtgctaaagatgggataattcctgag 806
Db 739 GTCTGTGTGTC---TGTTGATTACGTGAGGCAGTGTCTCTGACTCAGACAAATCCACCCC 795
QY 807 aacttcagtggtttcagtggttgcgggaaatcgagacagagagccttcattagttcaa 866
Db 796 AATTTCAGCCTCTTTGAAGTGGTCTGGAGATCGGAAACAGGACCTCGACGGGTGCAG 855
QY 867 acgcaggaaacatatgaactgctacaaatgctgtattagaactatt 913
Db 856 ACAGAGGAGCAGTACAGGTTTCCTGTACCAACACAGTGGCTTCAGCTATT 902

RESULT 4

US-08-036-210-10
; Sequence 10, Application US/08036210
; Patent No. 558233
; GENERAL INFORMATION:
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Moller, Karin B.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/036,210
; FILING DATE: 23-MAR-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2309 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1766
; OTHER INFORMATION: /note= "N=x-unknown nucleotide"

US-08-036-210-10

Query Match 5.5%; Score 130.6; DB 1; Length 2309;
Best Local Similarity 51.1%; Pred. No. 1.5e-26;
Matches 307; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

QY 173 tctacaactgtggctggagaatgccaagaatatcaagaaaaacagatatagaagatatattt 232
Db 184 TTCTTCAACTGATGCTGATCTGCCTTGGAAATAGACAAAAACCCGCTTCCCAACATAAA 243
QY 233 gccctatgattatagccgggttagaactatccctgataacacctgatgaggtatccagcta 292
Db 244 ACCATATAATAATAACAGAGTAAAGCTAGTAGCTACGCTAGTGTTCACAGGTTCCGATTGA 303
QY 293 catcaatgccaaactcattaaagggagtttatggaccaccaaggttatattgccaccaggg 352
Db 304 TATTAATGCCACCTATATTTCTGGTTATTTATGTCCAAATGAATTTATTTGCTACTCAAGG 363
QY 353 tctttatatacaacctcctggaactctggagatgattgggaatatagtgctcctat 412
Db 364 TCCACTACCAGGAACAGTTGGAGATTTTGGAGAAATGGTGTGGGAACCCAGACCAAAAC 423
QY 413 cattgttatggcatgcatgagatgataatgaaatgggaaagaaagtgtagcgcctactgggc 472
Db 424 ATTAGTAATGCTAACACACAGTGTTTTGAAAAAGACGAGATCAGATGCATCAGTATTGGCC 483
QY 473 tgagccaggagagatgcatggaatttgcccttctctgtatcctgtgaagctgaaaa 532
Db 484 AGAGGACAAACAGCCAGTTACTGTCTTTGGAGATATAGTGTATACAAAGCTAATGGAGGA 543
QY 533 aaggaaatcgtattataatcaggaactctaaagttaaagttcaatagtgaaactcgaac 592
Db 544 TGTTCAAATAGATTGGACTATCAGGATCTGAAAAATTGAAAGGCATGGGATTCGATGAC 603
QY 593 tatctaccagtttcatcaagaatggccagaccatgatgtaccttcattctatagaccc 652
Db 604 TGTTCACACAGTGTAACTTTACTGCCTGCCGACAGCATGGGTTCCTGAGAACAGCCGCC 663
QY 653 tattcttgagctcatctgggatgtagtctgttaccagaagagatgacagtgcttcccatag 712
Db 664 TCTAATTCACCTTTGTGAAGTTGGTTCGACAGAGCCAGGCACATGACACACACCTATGAT 723
QY 713 cattcactcagtgctggctgggaaggactggtgtatttggctattgttgattatcac 772
Db 724 TGTTCACCTCAGTGTGGAGTTGGAAGAACTGGAGTTTATTGCTCTCGACCATTTAAC 783
QY 773 a 773
Db 784 A 784

RESULT 5

US-08-449-609-10
; Sequence 10, Application US/08449609
; Patent No. 5952212
; GENERAL INFORMATION:
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Moller, Karin B.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

Db      724  TGTTCACTGCAGTCTGGAGTTGGGAAGAACTGGAGTTTTTATTGCTCTGGACCATTTAAC  783
Qy      773  a 773
      |
Db      784  A 784

RESULT 6
US-08-036-210-14
; Sequence 14, Application US/08036210
; Patent No. 5585233
; GENERAL INFORMATION:
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Moller, Karin B.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

:
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/036,210
: FILING DATE: 23-MAR-1993
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
:
: NAME: Misrock, S. Leslie
:
: REGISTRATION NUMBER: 18,972
: REFERENCE/DOCKET NUMBER: 7683-025
: TELECOMMUNICATION INFORMATION:
:
: TELEPHONE: 212-790-9090
: TELEFAX: 212-869-8864/9741
:
: TELEX: 66141 PENNIE
:
: INFORMATION FOR SEQ ID NO: 14:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 2692 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: unknown
: MOLECULE TYPE: cDNA
:
: FEATURE:
:
: NAME/KEY: misc_feature
: LOCATION: 2149
:
: OTHER INFORMATION: /note= "N-x=unknown nucleotide"
:
: US-08-036-210-14
:
:
:
: Query Match 5.5%; Score 130.6; DB 1; Length 2692;
: Best Local Similarity 51.1%; Pred. No. 1.6e-26;
: Matches 307; Conservative 0; Mismatches 294; Indels 0; Gaps
:
:
: QY 173 tctacaactgtgcgtgagaatgccagaatatcaagaaaacagataataagatatattt 232
: | | | | | | | | | | | | | | | | | | | | | | | |
: Db 567 TTCTTCAACTGATGCTGATCTGGAAATAGACCAAAACCGCTTCCCAACATATA 626
: | | | | | | | | | | | | | | | | | | | | | | | |
: QY 233 gccctatgattagccgggtagaactatccctgataaactctgatgaggattccagcta 292
: | | | | | | | | | | | | | | | | | | | | | | | |
: Db 627 ACCATATAATAAACAGAGATAAAGCTGATAGCTGACGCTAGTGTCTCCAGGTTCCGATTA 686
: | | | | | | | | | | | | | | | | | | | | | | | |
: QY 293 catcaatgccaaacttcattaagggaggttatgaccacaaggcttatatgccaccaccagg 352
: | | | | | | | | | | | | | | | | | | | | | | | |
: Db 687 TATTAATGCCAGCTATATTTCTGGTTATTTATGTCCAAAATGAATTTTGTCTACTCAAGG 746
: | | | | | | | | | | | | | | | | | | | | | | | |
: QY 353 tcccttatctacacacctcctgcactctcggagatgatttggqaataaagtcctctat 412
: | | | | | | | | | | | | | | | | | | | | | | | |

```

Query Match 5.5%; Score 130.6; DB 1; Length 2692;
Best Local Similarity 51.1%; Pred. No. 1.6e-26;
Matches 307; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

Db 747 TCCACTACCGAACAGTGGAGATTTTGGAGATGCTGTGGAAACACGAGCAAAAAC 806
Qy 413 catgttatgcatgcatgagatgaatggaaagataaaagtgtgagcctactgggc 472
Db 807 ATTAGTAATGCTAACACAGTGTGTGAAAAAGGACGATCAGATGCCATCATATTGGCC 866
Qy 473 tgagccaggagagatgcagctggaatttggcccttctctgtatctgtgaagctgaaaa 532
Db 867 AGAGCACACAGCCAGTACTGCTTTGGAGATATAGTATCAAAAGCTAATGGAGGA 926
Qy 533 aaggaaatctgattataatacagagacttaaaagttaagttaaatagtgaaactcgaa 592
Db 927 TGTTCAAATAGATTGACATATCAGGATCTGAAATTCAAAAGGATGGGATTCGATCAC 986
Qy 593 tatctaccagtttcatcaagaattggccagaccatgatgtacaccttcatctatagacc 652
Db 987 TGTTCGACAGTGAATCTTACTGCTGCCAGAGCATGGGTTCTTGAGAACACGCGCCC 1046
Qy 653 tattcttgagctcatctgggatgtacgtgttaccagaagatgacagtgctcccatatg 712
Db 1047 TCTAATTACATTGGAAGTTGGTTCGAGCAAGCAGGACACATGACACACCTATGAT 1106
Qy 713 cattcaatgcagctgctggctgtggaagagctgtgttatttctgtctattgtgtattac 772
Db 1107 TGTTCACCTGCAGTGTGGAGTTGGAAGAACTGGAGTTTTTTATTGCTCTGGACCATTTAAC 1166
Qy 773 a 773
Db 1167 A 1167

RESULT 7
US-08-449-609-14
; Sequence 14, Application US/08449609
; Patent No. 5952212
; GENERAL INFORMATION:
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Moller, Karin B.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE
; NUMBER OF INVENTION: PHOSPHATASE
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,609
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/036,210
; FILING DATE: 23-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2692 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2149
; OTHER INFORMATION: /note= "N=x-unknown nucleotide"
US-08-449-609-14

Query Match 5.5%; Score 130.6; DB 2; Length 2692;
Best Local Similarity 51.1%; Pred. No. 1.6e-26;
Matches 307; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

Qy 173 tcctacaactgtgctgagaatgccaagaataatcaagaaataagataagataatttt 232
Db 567 TTCTTCACATGATGCTGATCTGCCTTGGAAATAGAGCAAAAACCGCTTCCCAACATAAA 626
Qy 233 gccctatgattatagccgggtagaaactatccctgataaacctctgatgaggattccagcta 292
Db 627 ACCATATATATACACAGTAAAGCTGATAGCTACCGCTAGCTGTCCAGGTTCCGATTA 686
Qy 293 catcaatgccaaacttcattaaagggagtttatggaccccaagcttatattgcccacccaggg 352
Db 687 TATTAATGCCAGCTATATTTCTGTTTATTTATGTCCTGTCCTGTCCTGTCCTGTCCTGTC 746
Qy 353 tcccttatctacaacccctcctggacttctggagatgatttgggaataatagttccttat 412
Db 747 TCCACTACCGAACAGTGGAGATTTTGGAGATGCTGTGGAAACACGAGCAAAAAC 806
Qy 413 catgttatgcatgcatgagatgaatggaaagataaaagtgtgagcctactgggc 472
Db 807 ATTAGTAATGCTAACACAGTGTGTGAAAAAGGACGATCAGATGCCATCATATTGGCC 866
Qy 473 tgagccaggagatgcagctggaatttggcccttctctgtatctgtgaagctgaaaa 532
Db 867 AGAGCACACAAAGCCAGTACTGCTTTGGAGATATAGTATCAAAAGCTAATGGAGGA 926
Qy 533 aaggaaatctgattataatacagagacttaaaagttaagttaaatagtgaaactcgaa 592
Db 927 TGTTCAAATAGATTGACATATCAGGATCTGAAATTCAAAAGGATGGGATTCGATCAC 986
Qy 593 tatctaccagtttcatcaagaattggccagaccatgatgtacaccttcatctatagacc 652
Db 987 TGTTCGACAGTGAATCTTACTGCTGCCAGAGCATGGGTTCTTGAGAACACGCGCCC 1046
Qy 653 tattcttgagctcatctgggatgtacgtgttaccagaagatgacagtgctcccatatg 712
Db 1047 TCTAATTACATTGGAAGTTGGTTCGAGCAAGCAGGACACATGACACACCTATGAT 1106
Qy 713 cattcaatgcagctgctggctgtggaagagctgtgttatttctgtctattgtgtattac 772
Db 1107 TGTTCACCTGCAGTGTGGAGTTGGAAGAACTGGAGTTTTTTATTGCTCTGGACCATTTAAC 1166
Qy 773 a 773
Db 1167 A 1167

RESULT 8
US-08-036-210-21
; Sequence 21, Application US/08036210
; Patent No. 5585233
; GENERAL INFORMATION:
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Moller, Karin B.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE
; NUMBER OF INVENTION: PHOSPHATASE
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS

STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/036,210
FILING DATE: 23-MAR-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7683-025
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 3973 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc.feature
LOCATION: 3430
OTHER INFORMATION: /note= "N=x-unknown nucleotide"
US-08-036-210-21

Query Match 5.5%; Score 130.6; DB 1; Length 3973;
Best Local Similarity 51.1%; Pred. No. 2e-26;
Matches 307; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

QY 173 tctacaactgtggtgaggaatcccaagaatatcaagaaaaacagatatagaatatttt 232
DB 1848 TTCTTCAACTGATGCTGATCTGCTTGGTAATAGAGCAAAAACCGCTTCCCAACATATAA 1907

QY 233 gccctatgattagcgggtagaactatccctgataacacctgtagagattccagcta 292
DB 1908 ACATATAATTAACAGAGTAAGCTGATGCTGACGCTAGTGTTCACAGGTTCCGGATTA 1967

QY 293 cacaatgcacacttcattaaaggaggtttatggaccacaaaggcttatattgccaccagg 352
DB 1968 TATTAATGCCAGCTATATTTCTGTTATTTATGTCCAAATGAATTTATTGCTACTCAAG 2027

QY 353 tcttttatcacaacctctcgactctcgaggatgatttgggaatatagtgctccttat 412
DB 2028 TCCACTACCAAGGAACAGTTGGAGATTTTGGAGAATGGTGTGGGAAACACAGACAAAAC 2087

QY 413 cattgttagcatgcatgagatgagatgaatggaaagaaagtgagcgctactgggc 472
DB 2088 ATTAGTAATGCTAACACAGTGTTTGAAAAGGACGGATCAGATGCCATCAGTATTGGCC 2147

QY 473 tgagccaggagagatgcagctggaatttggcccttctctgtatccctgtgaagctgaaa 532
DB 2148 AGAGGACAACAGCCAGTTACTGCTTTGGAGATATAGTATTAACAAGCTAATGGAGGA 2207

QY 533 aaggaaatcgtattataataatcaggactctaaagttaagttaactgaactgaac 592
DB 2208 TGTTCAAATAGATTGGACTATCAGGGATCTGAAAATTTGAAAGGCATGGGGATTGCGATGAC 2267

QY 593 tatctaccagtttcattacaagaattggccagaccatgatgtacccttcattatagacc 652
DB 2268 TGTTCGACAGTAACTTTTACTGCTGGCCAGGACATGCGGGTTTCCCTGAGAACAGCGCCCC 2327

QY 653 tattcttgagctatctctgggatgtagctgtttaccagaagagatgacagtggttcccatatg 712

DB 2328 TCFAATTCACCTTTGTGAAGTTGGTTCGAGCAAGCAGGGCACATGACACACACCTATGAT 2387
QY 713 cattcaactgcagtgctgctgaggaactggtgtatttatttgctatttattgtattatatac 772
DB 2388 TGTTCACCTGAGTCTGCTGAGTTGGAAAGAACTGGAGTTTATTGCTCTGGACCATTTAAC 2447

QY 773 a 773
DB 2448 A 2448

RESULT 9
US-08-449-609-21
Sequence 21, Application US/08449609
Patent No. 5952212
GENERAL INFORMATION:
APPLICANT: Moller, Niels P.H.
APPLICANT: Moller, Karin B.
APPLICANT: Ullrich, Axel
TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE
PHOSPHATASE
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,609
FILING DATE: 24-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/036,210
FILING DATE: 23-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7683-025
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 3973 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc.feature
LOCATION: 3430
OTHER INFORMATION: /note= "N=x-unknown nucleotide"
US-08-449-609-21

Query Match 5.5%; Score 130.6; DB 2; Length 3973;
Best Local Similarity 51.1%; Pred. No. 2e-26;
Matches 307; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

QY 173 tctacaactgtggtgaggaatcccaagaatatcaagaaaaacagatatagaatatttt 232
DB 1848 TTCTTCAACTGATGCTGCTTGGTAATAGAGCAAAAACCGCTTCCCAACATATAA 1907

QY 233 gccctatgattagcgggtagaactatccctgataacacctgtagagattccagcta 292


```

; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Gallus domesticus
; STRAIN: D740
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 91..3802
US-08-850-917-1

Query Match
Best Local Similarity 4.9%; Score 115.4; DB 2; Length 3969;
Matches 336; Conservative 0; Mismatches 22; Indels 9; Gaps 1;

Qy 164 caaaacctatctcaaaccttggtggaagtccaagaatatcaagaaaaacagatataa 223
Db 1881 CACTAAATTTCCATGAGGAGGCCAAGAGGCCATTAATCAGACAAAAACCGTTACAT 1940

Qy 224 ggatatgttgcctatgattagccgggtagaaactatccctgataaaccctctgatgaga 283
Db 1941 TGATATTCTTCATATGATCATAACCGTGTGAGCTCTCTGAGATTCAGGAGACCCAGG 2000

Qy 284 ttccagctacatcaatgccaaacttcaatgaaggagtttatggaccgaaggttatatgc 343
Db 2001 ATCAGACTACATCAACGCAAGTTATATTGATGGCTTCAAAGAACCGAGAAAAATACATTGC 2060

Qy 344 caccagggctctttatctcaaacctctgactcttgaggatgattgggaatatag 403
Db 2061 TGCACAGGCCCAAGGATGAACACCGGATGATTCTCGGAGATGATCTGGGAACAGAA 2120

Qy 404 tgccttatcatctgttggcatgcatgagtgatgaaatggaaagaaaaagtgtgagcg 463
Db 2121 AGCAACAATTTATTGTCATGTTACTTCGCTGTGAGGAAGAAACAGGAACAATGTGCCCA 2180

Qy 464 ctactgggctgagccagagagatgagctggaatttggcccttctctgtatctgtga 523
Db 2181 GTACGGGCCATCAATGGGAGATGGCTCTGCAACATATGGGGACACATAACTGTGAAGATCAA 2240

Qy 524 agctgaaaaaagaaactgattatataatcagactctaaaagttaagttcaatagtga 583
Db 2241 CGAAGTAAATATGTCAGACTATATATTCAGAAACTGACACATFCACAAATGGAAGAGA 2300

Qy 584 aactgaaac-----tatctaccagtttctattacagaatttgcagaccatgatgt 634
Db 2301 AAGAACATCTGGAAGAGATGTCACCTACATTCAGTTCACCACTGCGCCAGACCATGGCGT 2360

Qy 635 accttcactatagaccctatctttagctcatctctgagatgacgtttgtaccgaagga 694
Db 2361 TCCCAGGATCCACATCTCTCTTCAAACTCCGACGACAGTGAATGCTCTCAGCAACTT 2420

Qy 695 tgacagtttcccatatgactcactgcagctgctggtgtggaaggactgtgtatttg 754
Db 2421 TTTAGTGGCCCAATAGTGGTTCATGTGAGTGCTGGAGTTGGGGCAGCTGGGACCTATAT 2480

Qy 755 tgcattgttgattacatagatgtgtgtcctaaagatgggataatctctggaacattcag 814
Db 2481 TGGAAATCAGCTATGTTGGAGGGCTGGATGCAGAGGGCAGAGTGGATGTTTATGGGTA 2540

Qy 815 tgttttcagttgatccggga 835
Db 2541 CGTTGTGAAGCTGCGCGGCA 2561
```

```

; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: 60/006,838
; FILING DATE: 16-NOV-1995
; FILING DATE: 30-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/006,838
; FILING DATE: 16-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Remenick, James
; REGISTRATION NUMBER: 36,902
; REFERENCE/DOCKET NUMBER: 0A146-0110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; TELEX:
; INFORMATION FOR SEQ ID NO: 74:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; US-08-594-031-74

Query Match
Best Local Similarity 4.5%; Score 106.2; DB 1; Length 247;
Matches 159; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

Qy 112 ttgcaatgaatttctgaagctgaaaaggcaatctaccaagtacaaggcagacaaaacct 171
Db 1 TCGCCCGGAGCTTCATGCGATTGAGAGATTGTCTACCAATATAGAACACAGAAAGATTT 60

Qy 172 atctacaactgtgctgagaaatgcccaagaatatcaagaaaaacagataaaggatatt 231
Db 61 ATCCACAGCCACTGGAGAAAAAGAGAAATGTTAAAAAGACACAGATATAAGGACATAC 120

Qy 232 tgcctatgattatagccgggtagaactatccctgataaaccctctgagagattccagct 291
Db 121 TGCCATTTGATCAGACGCGAGCTTAAGTTGACTTTGAAGACTCCCATCCCAAGATTTCAGATT 180

Qy 292 acatcaatgccaaactcattaaaggagtttatgacccaaggcttatattgcccaccagg 351
Db 181 ATATCAATGCAAAATTTTATTAAAGGTGTGTATGGCCCAAAAGCATATGTGCAACCCCAAG 240

Qy 352 gtccttt 358
Db 241 GGCCTTT 247
```

Search completed: May 26, 2002, 10:33:59
Job time: 12816 sec

```

RESULT 15
US-08-594-031-74
; Sequence 74, Application US/08594031
; Patent No. 5783182
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Timothy C.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2002, 06:59:45 ; Search time 5560.98 Seconds
(without alignments)
8865.869 Million cell updates/sec

Title: US-09-600-358A-3
Perfect score: 2356
Sequence: 1 tccctcaacctacttataga.....cctagtcggagcacatccgg 2356

Scoring table: IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.:

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description

1	2356	100.0	2356	6	AX020263	Sequence
2	2352.8	99.9	2356	9	AF001847	Homo sapi
3	2097	89.0	3058	6	AX020261	Sequence
4	2093.8	88.9	3058	9	AF001846	Homo sapi
5	2074.6	88.1	3580	6	AF149916	Sequence
6	1761	74.7	2877	9	AF150732	Homo sapi
7	1261	53.5	2734	10	MUSPROTYPH	M03888 Mouse prote
8	816.8	34.7	143618	9	AL137856	Human DNA
9	447.2	19.0	1817	9	BC017785	Homo sapi
10	378.2	16.1	3160	9	HUMTPPEST	M93425 Human prote
11	376.6	16.0	2938	9	HUMTPPGI	D13380 Human mRNA
12	376.6	16.0	2938	11	GI8337	X18337 SWISS1985 Er
13	370.2	15.7	2676	10	MMPTPPES	X86781 M.musculus
14	368.6	15.6	1608	10	RATRKPTP	D38072 Rat mRNA fo
15	368	15.6	2983	6	AX306312	Sequence
16	368	15.6	2983	10	MM19PTP	X63440 M.musculus
17	243.6	10.3	253	9	HSU69700	U69700 Human prote
18	218	9.3	1529	10	MMPTPHSC	U9853 Mus musculu
19	218	9.3	1559	10	MMU35124	U35124 Mus musculu
20	218	9.3	1568	10	MMU52523	U52523 Mus musculu
21	216.4	9.2	1529	6	AR154418	Sequence
22	216.4	9.2	1643	10	BC008512	Sequence
23	214.2	9.1	2198	10	RNU69673	U69673 Rattus norv
24	214.2	9.1	2226	6	AR095603	X79568 H.sapiens B
25	202.2	8.6	2810	9	HSBDP1	X79568 H.sapiens B
26	144.6	6.1	4315	9	HSLCAR	Y00638 Human mRNA
27	144.6	6.1	4597	9	HSLCA	Y00638 Human mRNA
28	139.6	5.9	4655	10	MUSLY5A	M14342 Mouse 200 k
29	139.6	5.9	5186	10	MUSLY5B	M14343 Mouse Ly-5
30	136.4	5.8	5007	10	MUSLY5AA	M92933 Mouse lymph
31	133	5.6	3541	10	RATLICAL	M10072 Rat leukocy
32	133	5.6	4027	10	RNLCAR	Y00065 Rat mRNA fo
33	130.6	5.5	2309	6	AR073853	Sequence
34	130.6	5.5	2309	6	I320377	I320377 Sequence 10
35	130.6	5.5	2692	6	AR073854	Sequence 14
36	130.6	5.5	2692	6	I32038	Sequence 14
37	130.6	5.5	3973	6	AR073855	Sequence 21
38	130.6	5.5	3973	6	I32039	Sequence 21
39	121	5.1	4920	5	AF173857	Xenopus 1
40	120	5.1	466	6	AR154432	Sequence
41	116.8	5.0	4756	9	HUMTPPS	D37781 Human mRNA
42	116.8	5.0	5117	6	AR109903	Sequence
43	116.8	5.0	5117	9	HSU10886	U10886 Human densi
44	116.2	4.9	4802	10	MMMRPTPU	X58287 Murine mR-P
45	115.6	4.9	6314	10	RNU40790	U40790 Rattus norv

ALIGNMENTS

RESULT 1	AX020263	Sequence 3 from Patent WO9936548.	2356 bp	DNA	linear	PAT 07-SEP-2000
AX020263	LOCUS	AX020263	Sequence 3 from Patent WO9936548.			
	DEFINITION	AX020263				
	ACCESSION	AX020263				
	VERSION	AX020263.1	GI:10044044			
	KEYWORDS					
	SOURCE	human.				
	ORGANISM	Homo sapiens				
		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
	REFERENCE	1 (bases 1 to 2356)				
	AUTHORS	Roifman,C.M.				
	TITLE	Human lymphoid protein tyrosine phosphatases				
	JOURNAL	Patent: WO 9936548-A 3 22-JUL-1999;				
		HSC RES DEV LP (CA); ROIFMAN CHAIM M (CA)				
	FEATURES	Location/Qualifiers				
	source	1..2356				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
	CDS	42..2120				
		/note="unnamed protein product"				
		/codon_start=1				

/protein_id="CAC07769.1"
/db_xref="GI:10044045"
/translation="MDREILQKLFDEAQSKKIKKEEFANFLKLRKSTVYKADKTV
PTVAENAKNKKYKIDILPDYDSVELSLITDSDSSYINANFIVGYPKAYIAT
QGPTSLDLDLFWIMWEXSVLIIVMACMEYEMGKKCKERYWAEPGEMQLFEGPFSVC
EAEKKRSIIIRLUKVPNSETRTIYQPHYKNPDPHVPSSIDILEILWIDVRCYQD
DSVPICHSACGCTGVICAIYDVTWMLLDGIIIPENFSVSLIREMRTQPSLVQT
QEOVELVYNAVLELFRKQMDVIRKHSGETSOAKHCIPKHNHTLQADSISPLPKSTT
KAAKMNQORIKMETKSSSFDFTSEISAKEELVLPKAKSSTSFDFLELANSFDPKNA
DITMKWQKAPPIVGEPLQKHQSLDGLSLLFEGCSNPKPVNAGRFNSKVPITPKS
TPFELIQOREKEVDKSNFSLQPHDSCFVEMQAKYMHVSAEYNSLIPDYSKH
QIRNASNKHDSALGVYSYIPIVNPENFSYPPSPSSKMSLDLPBKQDGPFPSS
LLPSSLSFYSHSLNSPLNTISLLNQESAVLATAPRIDDELPTPLPVRTPPE
SFIVVEEAGEFSPNPKVSLSSAVKVIKGLTSLEWGGTSEPKKFKDDSVILRPSKSVKLRS
PKSGKNFSWL"

BASE COUNT		750 a	466 c	466 g	674 t
ORIGIN					
Query Match 100.0%; Score 2356; DB 6; Length 2356;					
Best Local Similarity 100.0%; Pred. No. 0;					
Matches 2356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	tcctcctaactctatagactattttctctctgcagcatggaccacaagagaattc	60		
Db	1	TCCTCAACCTACTTATAGACTATTTTCTGCTGCGAGCATGGACCAAGAGAAATTC	60		
QY	61	tgcagaagttctgcatgagccgcaaaagcaagaataattactaaagagatttgcgaatg	120		
Db	61	TGCAGAAGTTCTCTGGATGAGGCCCAAGCAAGAAATTTACTAAAGAGGATTTGGCAATG	120		
QY	121	aattctgaagctgaaaagcgaactctaccaagtacaaagcagacagacaaacctatcctacaa	180		
Db	121	AATTTCTGAAGCTGAAAGGCAATCTACCAAGTACAAGCGACGACAAACCTATCTACAA	180		
QY	181	ctgtggctgagaaatgccagaatatcaagaaaacagatatagaagatatatttgccctatg	240		
Db	181	CTGTGGCTGAGAATGCCAAGATATACAGAAAACAGATATAGGATATATTTGCCCTATG	240		
QY	241	attatagccggtagaactatccctgataacctctgcatgagattccagctacatcaatg	300		
Db	241	ATTATAGCCGGGTAGAACCTATCCCTGATAACTCTGATGAGGATTCAGACTACATCAATG	300		
QY	301	ccaaacttcataagggagtttatggaccacaagcttatattgccaccagggtccctttat	360		
Db	301	CCAACCTTCAATTAAGGAGTTTATGACCCCAAGCCTTATATGCCACCCAGGTCCTTTAT	360		
QY	361	ctacaacctctgactctctggagatgatttgggaatatagtgctccttatcattgtta	420		
Db	361	CTACAACCTCTCTGGACTTCTGGAGGATGATTTGGGAATATAGTGTCTTATCATTTGTTA	420		
QY	421	tggcatgcatgagatgaaatggaaaagaaaagtgtagcgctactgtggtgagccag	480		
Db	421	TGGCATGCTAGGATGATGAATGGGAAAAGAAAAGTGTGAGCGCTACTGGCTGAGCCAG	480		
QY	481	gagagatgagctggaaatttggcccttctctgtatccctgtgaagctgaaaaagaaat	540		
Db	481	GAGAGATGAGCTGGAATTTGGCCCTTTCTCTGTATCTCTGTGAAGCTGAAAAAGAAAT	540		
QY	541	ctgattataatacaggactctaaaagttaagtccaatagtgtaaacctgaactatatacc	600		
Db	541	CTGATTATATAATCAGGACTCTAAAAGTTAAGTTCAATAGTGAACCTCGAACTATCTACC	600		
QY	601	agtttcattacaagaattggccagaccatgatgtaccttcctctatagaccctattcttg	660		
Db	601	AGTTTCATTACAAGAAATTTGGCAGACCATGATGATACCTTCATCTATAGACCTTATTTCTTG	660		
QY	661	agctcatctggatgtacgttgttaaccaagagatgacagtggtcccatatgcatcact	720		
Db	661	AGCTCATCTGGGATGTACGTGTGTACCAAGAGATGACAGTGTTCCTCATATATGCAITCACT	720		
QY	721	gaagtctggctgtgaaaggaactggttatttggctatttggattataataatgatg	780		
Db	721	TGAATTTCTCCACCAATATTTCTCTACTATTGAACCAGGAGTCAGCTGTACTAGCAACTG	780		

Db	721	GCAGTGCTGGCTGTGGAAAGGACTGGTGTATTATTGTGCTATTGTGTATTATACATGATGT	780
QY	781	tgtctaaagatggataaattcctcgagaacttcagtgcttttcagtttgatccgggaatgc	840
Db	781	TGCTAAAGATGGGATAAATCTTGAGAACTTCAGTGTCTTTTTCAGTTTGATCCGGGAATGC	840
QY	841	ggacacagagggctctcattagttgtaaacccagagaacaatatgaactggctcacaatgctg	900
Db	841	GGACACAGAGGGCTTCATTAGTTCAAACCCAGAGAACAAATGAACCTGGTCTCAACATGCTG	900
QY	901	tattagaactatttaagagacagatggatgttatcagagataaaacattctggaacagaga	960
Db	901	TATTAGAACTATTTAAGACACAGATGGATGTTATCATCAGAGATAAACATTTCTGGAACAG	960
QY	961	gtcagaagcaagcattgtattcctcgagaaaaatcacactctcaagcagactcttattctc	1020
Db	961	GTCAAGCAAGCATGTATTCTCTGAGAAAATCACACTCTCCNAGCAGACTCTTATTCTC	1020
QY	1021	ctaatttaccaaaaagtaccacaaagcagcaaaaatgatgaaccacaaagacaaaaa	1080
Db	1021	CTAATTTACCAAAAAGTACCACAAAAGCAGCAAAAATGATGAACCAACAAGGACAAAAA	1080
QY	1081	tggaaatcaagaatcttctctcttggacttttaggactcttgaataagtgcacaaagaag	1140
Db	1081	TGGAATCAAGAATCTTCTCTCTTTGACTTTTAGGACTTCTGAAATAAGTGCACAAAAG	1140
QY	1141	agctagtttgcaccctgctaatacagcactcttctgactctctggagctaaattaca	1200
Db	1141	AGCTAGTTTGCACCTGCTNAATCAAGCACTTCTTTTTCGACTTCTGAGGCTTAATTACA	1200
QY	1201	gtttgacaaaaatgctgacacaaacctgaaatggcagacaaaggcatttcccaatagttg	1260
Db	1201	GTTTTCACAAAAATGCTGACACAAACCATGAATGGCAGACAAAAGGCATTTCCAATAGTTG	1260
QY	1261	ggagacctctcagaagcatcaaaagtttgagttgggctctcttggtaggggagttg	1320
Db	1261	GGAGCCTCTTCAGAAAGCATCAAAGTTTGGATTTGGGCTCTCTTTTGTGTAGGGGATGT	1320
QY	1321	ctaattctaactgtaaatgcagcaggaagatatatttaattcaaaagggtgccataaacac	1380
Db	1321	CYAACTTAACCTGTAAATGCAGCAGGAAGATATTTTAATTCAAAGGTGCCAATAACAC	1380
QY	1381	ggacacaaatcaactcctttgaaatgatagcagcagagagaaacaaaggaggtggacagca	1440
Db	1381	GGACCAAAATCAACTCCTTTTGAATTCATACGACAGAGAGAAAACAAGGAGGTGGACAGCA	1440
QY	1441	aggaacactttcttatttgggaatctcaaccacatattctgtttttgtagaatgcagg	1500
Db	1441	AGGAAACCTTTTCTTATTGGAAATCTCAACCACATGATTCTTTGTTGTAGAGATGCAGG	1500
QY	1501	ctcaaaaagtaatgcagtgttcttcagcagaactgaattattcaactgccaatagactcta	1560
Db	1501	CTCAAAAAGTAAATGCATGTTTCTTCAGCAGAACTGAATTAATTCACCTGCCATATGACTCTA	1560
QY	1561	aaacacaaatacgttaatgcctctaaatgtaaagcaccatgactctagtcctcttgggtat	1620
Db	1561	AACACCAAAATACGTAATGCTCTTAATGTAAGCACCATGACTAGTGTCTTCTTGGTGTAT	1620
QY	1621	attcttaacataaccttagtggaaaaatcccttatttttcatcatcgccccaagtggtacca	1680
Db	1621	ATTCTTACATACCTTTAGTGGAAAATCCCTTATTTCATCATGGCCTCCCAAGTGGTACCA	1680
QY	1681	gttctaagatgtctcttgatttacctgagaagcaagatggaactgttttccctcttctc	1740
Db	1681	GTTCTAAGATGCTCTTGTGATTACCTTGAGAAAGCAAGATGGAACCTGTTTTCTTCTTCTC	1740
QY	1741	tgttgcaacatctctcatcctctcttcttcttattacaattccacatagttcttattcac	1800
Db	1741	TGTTGCCAACATCCTCTACATCCTCTCTTCTTATTACAAATTCACATAGTCTCTTTATCAC	1800
QY	1801	tgaattctcaaccaataatttctcactatttgaacagagatgcagctgttactagcaactg	1860
Db	1801	TGAATTTCTCCACCAATATTTCTCTACTATTGAACCAGGAGTCAGCTGTACTAGCAACTG	1860

Qy 1861 ctccaagatagatgataaattccccctccacttccctgtagcacacctaataattta 1920
Db 1861 CTCCAAGATAGATGATAAATCCCCCTCCACTCCCTGTTAGGCACACTGAATCATTTA 1920
Qy 1921 ttgtggttgaggagctggagaattctcaccataattgttcccaaatcttattccctcagctg 1980
Db 1921 TTGTGGTTGAGGAGCTGGAGAATTTCTCACCATAATGTTCCCAAAATCCTTATCCTCAGCTG 1980
Qy 1981 tgaaggttaaaattggaacatcactggaattgggttggaacatctgaacccaagaatttg 2040
Db 1981 TGAAGGTAAAAATTGGAACATCACTGGAAATGGGTGGGAACATCTGAACCAAGAAATTTG 2040
Qy 2041 atgactctgtgatacttagacccaagcagagtgtaaaactccgaagtccttaaatcagcta 2100
Db 2041 ATGACTCTGTGATCTTAGACCAAGCAGAGTGTAATACTCCGAAGTCTTAATCAGGTA 2100
Qy 2101 aaaatttctctgtgcttagatgacatttagccctaaagattgggaagaattggttgtaag 2160
Db 2101 AAAATTTCTCTTGGCTTTAGATGACATTTAGCCCTAAGATTGGAAGAAATGTTCTGTTAAG 2160
Qy 2161 tttagagtaattcacttcaggaagttagttggttcccaataatagcttccagattcattg 2220
Db 2161 TTTAGATTAATTCACCTCAGGAAGTTACTTGGTTCCCAATAATAGCTTCCAGTATTCAATG 2220
Qy 2221 atttattctgcttccagactagaaatttctgaagagtcattgggggaagctaggc 2280
Db 2221 ATTTATTTCTGGCTTCCCACTAGAAAATTTTGTAAAGAGTCATGGGGGAAGCTAGGGC 2280
Qy 2281 taaccagaaaaataaaataatgggataaaaaatcgaaactactgttttcccccta 2340
Db 2281 TAACCAGAAAAATAAATAATGGGATAAAAAATCGGAACACTACTGTTTCCCCCTA 2340
Qy 2341 gtcggagacatccgg 2356
Db 2341 GTCGGAGCACATCCGG 2356

RESULT 2
AF001847 AF001847 2356 bp mRNA linear PRI 01-OCT-2001
DEFINITION Homo sapiens lymphoid phosphatase Lyp2 mRNA, complete cds.
ACCESSION AF001847
VERSION AF001847.1 GI:4100633
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Cohen, S., Dadi, H., Shaoul, E., Sharfe, N. and Roifman, C.M.
Cloning and characterization of a lymphoid-specific, inducible
human protein tyrosine phosphatase, Lyp
Blood 93 (6), 2013-2024 (1999)
99168989
PUBMED 10068674
REFERENCE 2 (bases 1 to 2356)
Roifman, C.M.
Direct Submission
Submitted (30-APR-1997) Immunology and Allergy, The Hospital For
Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8,
Canada
FEATURES
source Location/Qualifiers
1..2356
/organism="Homo sapiens"
/db_xref="taxon:9606"
42..2120
/note="splice variant of lymphoid phosphatase Lyp1"
/codon_start=1
/product="lymphoid phosphatase Lyp2"
/protein_id="AAD00905.1"
/db_xref="GI:4100634"
/translation="MDQREILQKFLDEAQSQKITKEEFANFLKLKRSQTKYKADKTY

PTTVAENAKNKKRKYDILPYDYSRVLSLITSDSESSYINANFKVYGPXAYIAT
QGSLSTLLDFWIMWEYSVLIVMACMEYEMGKKCKERYWAPGEQMLEGPEPSVSC
EAKRKS DYILIRLKYKFNSETIYQFHYKNPDPDHPSSIDPISLIREMRTQPSLVGT
DEVPICHSAGCGRTGVCIAVDYTWMLKDGILPENFSYSLIREMRTQPSLVGT
QEOYELVYNAVLELFKRMQDVIIRKHSGETSQAKHCLPEKNHTLQADSYSPNLPKSTT
KAAMNQORTKMEIKESSEDFRTSISAKELVLHPAKSSTSFDFLELNYSPDKNA
DTFMKQTKAPPVIGPELQKSLDGLSLLFEGCSNKPVNAAGRYFNSKVPITRTKS
TPFELIOQRETKEVDKENFYLESDPHSCFVMOAQKVHVSASALNTSLPYDSKH
OIRNASNVKHHDSALGVYSYIPLVNPYFSSMPPSCTSSKMSLDLPEKODGTVFPSS
LLPSSLSLFSYNSHDSLSLNSPTNISLUNQESAVLATAPRIDDDIIPPLPVRTPE
SFIVVEEAGFSPNPKVSLSSAVKVIKGTISLEWGGTSEPKKFDSDSVILRPSKSVKLRS
PKSGKNFSWL"

BASE COUNT 750 a 466 c 466 g 674 t
ORIGIN

Query Match 99.9%; Score 2352.8; DB 9; Length 2356;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2354; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 tccctcaacctacttatagactattttcttgcctgcagcatggaccacaaagagaaattc 60
Db 1 TCCCTCAACCTACTATAGACTATTTTCTGCTCTGCAGCATGGACCAAGAGAAATTC 60
Qy 61 tgcagaagttcctgtagagcccaagcaagaaattactaaagagaggtttgcctaag 120
Db 61 TGCAGAAGTTCTCTGATGAGGCCCAAGCAAGAAAATTTACTTAAGAGGAGTTTGCCTAATG 120
Qy 121 aattctgaaagctgaaagggcaattctaccaagtaacaagggcagacaaacctatctcaaa 180
Db 121 AATTCTGAAGCTGAAAAGGCAATCTACCAAGTAGTAAGGCAAGACAAACCTATCTCTACAA 180
Qy 181 ctgtggtgagaatgccaaagataatacaagaaacagataataagatatattgcctctag 240
Db 181 CTGTGGCTGAGAATGCCAAGAAATATCAAGAAAAACAGATATAAGGATATTTTGCCTCATG 240
Qy 241 attatagccgggtagaactatccctgataacctctgagagattccagctacatcaatg 300
Db 241 ATTATAGCCGGGTAGAACTATCCCTGATAACCTCTGATGAGGATTCAGAGTACATCAATG 300
Qy 301 ccaacttcattaaagggagttattggaccacagggttatttgcacccagggttccttat 360
Db 301 CCAACTTCATTAAAGGGAGTTTATGGACCCCAAGGCTTATATTGCCACCCAGGCTCCTTAT 360
Qy 361 ctacaacctctctgactctgagagatgattgggaataagtgctcttactatctgta 420
Db 361 CTACAACCTCTCTGACTCTCGAGGAGATTTGGGAATATAGTGTCTCTTATCATTTGTTA 420
Qy 421 tggcatgcatgagtgatgaaatgggaagaaagtgtagcgctactgggctgagccag 480
Db 421 TGGCATGCTGAGTATGAATGGGAAGAAAAGTGTGAGCGCTACTGGCTGAGCCAG 480
Qy 481 gagagatgcatggaatttggcccttctctctgatactctgtagagctgaaaaaggaat 540
Db 481 GAGAGATGCTGCAATTTGGCCCTTCTCTGTATCTCTGTGAAGCTGAAAAGGAAT 540
Qy 541 ctgattataatcaggactctaaagttaagttaagttaagttaagttaagttaagtta 600
Db 541 CTGATTATATAATCAGGACTCTAAAAGTTAAGTTCAATAGTAGAAACTCGAACTATCTAC 600
Qy 601 agtttcattacaagaattggccagaccatgatgtacctctctctatagaccctattctg 660
Db 601 AGTTTCATTACAAGAAATTTGGCCAGACCATGATGTACCTTCATCTATAGACCTTATCTTG 660
Qy 661 agtcatctgggagtacgtttgttaccagaagagatgacagtggttcccatatgcattact 720
Db 661 AGCTCATCTGGGATGCTAGCTTTGTACCAAGAGATGACAGTGTTCCTCCATATGCAATCACT 720
Qy 721 gcagtgtgctgtggaagactggttatttctgtctattgttattgtattgattacatgagtg 780
Db 721 GCAGTGTGCTGTGGGAAGGACTGGTGTATTGTGTGCTATTGTGTGATTATACATGGAGTGT 780
Qy 781 tgcataaagatgggataattcctcagaaactcagtggttttcagtttgatccgggaaatgc 840

Db 781 TGTCTAAAGATGGGAATAATTCCTGAGAACCTTCAGTCTTTTCAGTTTCATCCGGGAATGC 840
QY 841 ggcacagagccttcattagttcaaacacagagaacaatatgaactggctcaaatgcctg 900
Db 841 GGCACAGAGGCCTTCATTAGTTTCAACCGCAGGAACAATGAACCTGGTCTCAATGCTG 900
QY 901 tattagaactatttaagagacagatggatgttatcagagataaacaattctcggacagaga 960
Db 901 TATTAGAACCTATTTAAGAGACAGATGGATGTATTATCAGAGATAAACATTCCTGGACACAGAGA 960
QY 961 gtcagcaaaagcattgtattccttgagaaaaatcacactctccaaagcagactcttattctc 1020
Db 961 GTCAAGCAAAAGCATTTGATTCTCTGAGAAAATCACACTCTCCAGCAGACCTTTATTCTC 1020
QY 1021 ctaattaccaaaaagtaccacaaaagcagcaaaaaatgatgaaccacaaagacaaaaa 1080
Db 1021 CTAATTTACCAAAAGTAGTACCACAAAAGCAGCAAAAATGATGAACCAACAAAGACAAAAA 1080
QY 1081 tggaaatcaagaactctctctcttgacttttagacttttagactttcgaataagtgcacaaagaag 1140
Db 1081 TGGAAATCAAGAATCTTCTCTCTTTGACTTTAGGACTTCTGAAATAGTGCACAAAGAG 1140
QY 1141 agctagtttgcacctcgttaatacgaactctctcttgaacttctcggagctaaataca 1200
Db 1141 AGCTAGTTTTCACCTGCTTAATCAAGCACTTCTTTTGACTTTCTGGAGCTAAATTACA 1200
QY 1201 gttttgacaaaaatgctgcacacacacatgaatggcagacaaagcattcccaatagt 1260
Db 1201 GTTTTGCACAAAATGCTGCACACACCATGAAATGGCAGACAAAGGCATTTCCCAATAGTTG 1260
QY 1261 gggagcctcttcagaagcatcaaatgttgattggctctctcttggtagggagtt 1320
Db 1261 GGGAGCCTCTTCAGAGCATCAAGTTTGGATTTGGCTCTCTTTGTTGAGGGATGTT 1320
QY 1321 ctaattcctaaactgttaaatgcagcgggaagataattttaattcaaaagggtcccaataaac 1380
Db 1321 CTAATTCTAAACCTGTAAATGSCAGAGGAAGATATTTTAAATCAAAAGGTGCCAATAACAC 1380
QY 1381 ggacaaaatcaactcctcttgaattgatcacagacagagaacaaagaggtggacagca 1440
Db 1381 GGACCAAAATCACTCTTTGAAATGATACAGCAGAGAGAGAAACCAAGAGGTGCACAGCA 1440
QY 1441 aggaactttcttatttggaaatctcaacacacatgattctgtttttagagatgcagg 1500
Db 1441 AGGAAAACCTTTCTTATTGGAATCTCAACACACATGATCTTCTTTTGTAGAGATGCAGG 1500
QY 1501 ctcaaaaagtaatgcattgttcttcagcagaactgaattattcaactgcacatgaactcta 1560
Db 1501 CTCAAAAAGTAATGCATGTTTCTTCAGCAGAACTGAATATTACCTGCCATATCACTCTA 1560
QY 1561 aacacaaaatcacgaatgcctctaagttaagcaccacatgaactgaactgctcttgggtgat 1620
Db 1561 AACACCAAAATACGTAATGCCTCTTAATGTAAGCACCATGACTCTAGTGCCTCTGGGTAT 1620
QY 1621 attctacacactcttagtggaaaaatccttatttttcaatgcctcccaagtggtacca 1680
Db 1621 ATCTTTACATACCTTTAGTGGAAAATCCTTATTTTTCATCATGSCCTCCCAAGGTGACCA 1680
QY 1681 gttctaagatgctcttatttaccctgagaagcaaatggaactgttttccctcttctc 1740
Db 1681 GTTCTAAGATGCTCTTGTATTTACCTGAGAAGCAAGATGAAGTGTATTTTCTCTCTCTC 1740
QY 1741 tgttgcaacatctctcacatcctctctcttattacaaattcacatgattctttatcac 1800
Db 1741 TGTGCGCAACATCTCTACATCCTCTTCTTATTATCAATTCACATGATCTTTATCAC 1800
QY 1801 tgaattctcaacaaaatatttctcactattgaaccaggagtagcagctgtactagcaactg 1860
Db 1801 TGAATTTCTCAACCAATATTTCCTCACTATTGAACCAGGAGTCAGCTGTACTAGCAACTG 1860
QY 1861 ctccaagatagatgatgaaatccccctccactctctgtcacgacacactgaaatcattta 1920
|||||

Db 1861 CTCAAGAGTAGATGATGAATCCCCCTCCACTCTCTACGACACACCTGAATCATTTA 1920
QY 1921 ttgtgttgaggaagctggagaattctcaccaaattctccaaattcttatcctcagctg 1980
Db 1921 TTGTGTGTGAGGAGCTGGAGAAATTCACCAAAATGTTCCCAAAATCTTATCTCAGCTG 1980
QY 1981 tgaagtaaaaaattggaacatcacactggaatgggtggaacatctgaaccaaagaaatttg 2040
Db 1981 TGAAGCTAAAAATTGGAACATCATCTGGAATGGGTGGAACATCTGAACCAAGAAATTTG 2040
QY 2041 atgactctgtgatacttagacaagaagaagtgataaaaactccgaagctcctaaatcaggta 2100
Db 2041 ATGACTCTGTGATACCTTAGACCAAGCAAGAGTGTAAAACTCCGAGTCTCTAAATCAGGTA 2100
QY 2101 aaaaattctctggcttttagatgacatttagccctaaagattggaagaatgggttcgttaag 2160
Db 2101 AAAATTTCTCTGGCTTTAGATGACATTTAGCCCTTAGATTGGAGAAATGGTTCGTTAAG 2160
QY 2161 tttagagtaattcaacttcagggaattacttggttcccataaatagcttccagttattcattg 2220
Db 2161 TTTTAGAGTAATTCACCTTCAGGAAGTTTACTGGTTCCCATAAATAGCTTCCAGTATTCTATTG 2220
QY 2221 atttattctggcttccacagactagaaaattttgtaaaagatcatcgtggggaagcctagggc 2280
Db 2221 ATTTATTTCTGGCTTTCCCGAGACTAGAAAATTTTGTAAAGAGTCATGGGGAACTAGGGC 2280
QY 2281 taaccgaaaaataaaaaataatgggataaaaaatcggaactactgtttttccccccta 2340
Db 2281 TAACCGAAAAATAAATAAATAAATGGGATAAANAATCGGAACACTACTGTTTCCCCCTA 2340
QY 2341 gtccggagcaatccgg 2356
Db 2341 GTCCGAGCACATCCGG 2356
RESULT 3
AX020261
LOCUS 3058 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 1 from Patent WO9936548.
ACCESSION AX020261
VERSION AX020261.1 GI:10044042
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3058)
AUTHORS Roifman,C.M.
TITLE Human lymphoid protein tyrosine phosphatases
JOURNAL Patent: WO 9936548-A 1 22-JUL-1999;
HSC RES DEV LP (CA); ROIFMAN CHAIM M (CA)
FEATURES
source Location/Qualifiers
1..3058
/organism="Homo sapiens"
/db_xref="taxon:9606"
42..2468
/note="unnamed protein product"
/codon_start=1
/protein_id="GI:10044043"
/db_xref="GI:10044043"
/translation="MDQREILQKFLDEAQSKKITKEEFANEFLLKLRQSTKVKYKADKTY
PTTVAENAKNIKKNRYKDIPLYDSRVELSLITSDSSYINANFKVGYGPKAYIAT
QGPLSTLLDFWRMTWEXSVLIIVMACMEYEMGKKCYWAEPEGMQLFEPFVSKH
EAEKRSKYIIRTLKVKNSERTTYQHYKNWPHDVPSSIDPILLELWVRCYQKH
DSVPICIHCSACGGTVCATVDYTMKDKGIIIPENFSVFLIREMTORPSLVQT
QOQYELVNAVLELFPKQMDVIRDKHSGTESQAKHICPEKNHTLOADSISNLPKSTT
KAAKMNQORTKMEIKESSFDFRTSEISAKEELVLHPAKSSTSFDFLNLNFSFKNA
DTMTKWOTKAFPIVGEPLQKHQSLDGLSLLFEGCSNKSVPVNAAGRYFNKVPITRKS
TFPELITQRETKEDSKENFSYLESQPHDSCEVEMQAQKVMHVSAAELNYSLPYDSKH
QJRNASNKVKHDSALGVYSIPLVENPYFSSWPPSGTSSKMSLDLPEKODGTVPSS
LLPTSTSLFSYVNSHSLNSLNSPTNISLLNOESAVLATAPRIDDEIPPLVPRTPE
STIVVEEAGFSPNPVKSLSSAVKVIKGTSLWEGGTSEPKFDDSVILRPSKVLRS
PKSELHQDRSPPPPLPPLERTLESFFLADEDCMQAQSITETYSITPDIMENSISKQTL

KTPGKSFTRSKSLKILRNMKKKSICSNPPNKPAPESVQSNSSSFLNFGFANRFSKPKG
PRNPPTWNI"

BASE COUNT 1015 a 592 c 553 g 898 t
ORIGIN

Query Match 89.0%; Score 2097; DB 6; Length 3058;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tccctcaactactatagactatttttctgtctgcagcatggaccacaagagaaattc 60
Db 1 TCCTCAACTACTATAGACTATTTTCTGTCTGCAGCATGGACCAAGAGAAATTC 60
Qy 61 tgcagaagttccctggatggcccaagcaagaaaaattactaaagagaggttgcgaatg 120
Db 61 TGCAGAAGTTCCCTGGATGAGGCCCAAGCAAGAAANATTAAGAGGAGATTGCCAATG 120
Qy 121 aattctgaagctgaaaaaggaactaccgaagtacaaaggcagacaaacccatctctacaa 180
Db 121 AATTTCTGAAGCTGAAAAGGCAATCTACCAAGTACAAGGCAGACAAAACCTATCTTACAA 180
Qy 181 ctgtggctgagaatgccaaagaatacaagaaaaaacagataaaggataattttgccctatg 240
Db 181 CTGTGGCTGAGAATGCCAAGAATATCAAGAAAAACAGATATAAGGATATTTTGCCCTATG 240
Qy 241 attatagccgggtagaaactccctgataacccctctgatgaggttcagctacatcaatg 300
Db 241 ATTATAGCCGGGTAGAACTATCCCTGNATAACCTCTGATGAGGATTCGAGCTACATCAATG 300
Qy 301 ccaacttcattaaggaggtttatggaccacaaggcttatattgccaccacagggctccttat 360
Db 301 CCAACTTCATTAAAGGGAGTTTATGGACCCCAAGGCTTATATTGCCACCACGGGCTTTAT 360
Qy 361 ctacaaacctctcggaactcttgaggatgatttgggaataatagtctccttatcatgtta 420
Db 361 CTACAAACCCCTCGGACTCTCGAGGATGATTTGGGAATATAGTGTCTTATCATTTGTTA 420
Qy 421 tggcatgctggaatgaaatgggaagaaaaagttgagcgtactggcgtgagccag 480
Db 421 TGGCATGCTGGAGTGAATGGGAAGAAAAAGTGTGAGCGCTACTGGGCTGAGCCAG 480
Qy 481 gagagatgcagctggaatttggcccttctctgtatctctgtaagctgaaaaagaaat 540
Db 481 GAGAGATGCAGCTGGAAATTTGGCCCTTCTCTGTATCTCTGAAGCTGAAAAAGGAAT 540
Qy 541 ctgattataaaccaggactctaaaagttaagtccaatagtgaaactcgaactatctacc 600
Db 541 CTGATTATATAACAGGACTCTAAAAGTTAAGTTCAATAGTGAACACTCGAACTATCTACC 600
Qy 601 agtttcaatcaagaattggccagacatgatgtaccttcatctatagacccttatcttg 660
Db 601 AGTTTCAATACAGAAATGGCCAGACCATGATGATACCTTCTATAGACCCCTATTCTTG 660
Qy 661 agctcatctgggtgacgttgttaccagaaggatgacaggtgtcccatatgattcaact 720
Db 661 AGCTCATCTGGGATGTACGTTGTACCAAGAGGATGACAGTGTCCCATATGATTCACCT 720
Qy 721 gcagtgctggctggaagacttgggttatttggctatgttgattatcacatgagatg 780
Db 721 GCAGTGTGCTGTGGAAAGACTGGGTGTTATTTGTGCTATTGTTGATTATACATGGATGT 780
Qy 781 tgcataaagatggataattccctgaaactcagttcttcaagtttgcagtcggaagatgc 840
Db 781 TGCTAAAGATGGGATAATTCCTTGAGAACTTCAGTGTGTTTTCAGTTTGCATCCGGGAATGC 840
Qy 841 ggacacagagcccttcattatgttcaacgcaggaaacaatatgaactgggtctacaatgctg 900
Db 841 GGACACAGAGCCCTTCATTAGTTCAACGCAGGAAACAATATGAACTGGTCTACATGCTG 900
Qy 901 tattagaactatttaagagacagatggatgttatacagagataaacattctctggacagaga 960
Db 901 TATTAGAACATATTAAAGAGACAGATGGATGTTTATCAGAGATAAACATCTCTGGAACAGAGA 960

Qy 961 gtcaagcaagcattgtattctctgagaaaaatcacactctcccaagcagactcttattctc 1020
Db 961 CTCAGCAAGACCATTTGATTCCTTGAGAAAATCACACTCTCCAGCAGACTCTTATTCTC 1020
Qy 1021 ctaattaccaaaaagtagccacaaaaagcagcaaaaaatgatgaaccaacaagaagacaaaa 1080
Db 1021 CTAAATTTACCAAAAAGTAGCCACAAAAGCAGCAAAAATGATGAACCAACAAGGACAAAAA 1080
Qy 1081 tggaaatcaagaatcttctctctttagacttttaggacttctgaaataaagtgcaaaagaag 1140
Db 1081 TGGAAATCAAGAATCTTCTCTTGTGACTTTAGGACTTCTGAAATAAAGTGCAAAAGAAG 1140
Qy 1141 agctagtttgcaccctgtataatcaagcacttcttcttcttcttcttcttcttcttcttct 1200
Db 1141 AGCTAGTTTGCACCCCTGCTAAATCAAGCACTTCTTTTGTGACTTCTGGAGCTAAATTA 1200
Qy 1201 gttttgacaaaaatgctgacacaaacccatgaaatggcagacaaaggcatttcccaatagt 1260
Db 1201 GTTTTGACAAAAATGCTGACACAACCATGAATGGCAGACAAAAGGCAATTTCCAATAGTTG 1260
Qy 1261 gggagcctcttcagaagcatcaaaagtttggatttgggctctcttcttcttcttcttcttct 1320
Db 1261 GGGAGCCTCTTCAGAAAGCATCAAAGTTTGGATTTGGGCTCTCTTTTGTGAGGGATGTT 1320
Qy 1321 ctaattctaaacctgtaaatgcagcaggaagatattttaaattcaaaaggtgccataaac 1380
Db 1321 CTAAATTTCAAACTGTAAATGCAGCAGGAAGATATTTTAATTCAAAGGTGCCAATTAAC 1380
Qy 1381 ggacaaaaatcaactccttttgaattgatacagcagagagaaaaagggaggtggacagca 1440
Db 1381 GGACAAAATCAACTCCTCTTTTGAATTTGATACAGCAGAGAGAAACCAAGGAGGTGGACAG 1440
Qy 1441 agaaaaacttttcttatttgggaattcacaaccacatgattcttcttcttcttcttcttct 1500
Db 1441 AGAAAACCTTTTCTTATTGGAAATCTCAACCACATGATTTCTTTTGTAGATGCGAGG 1500
Qy 1501 ctcaaaagcaatgcatgttcttcagcagaactgaattattcactgcacatatgactcta 1560
Db 1501 CTCANAAAGTAATGCATGTTTCTTCAGCAGAACTGAATTAATTCACCTGCCATATGACTCT 1560
Qy 1561 aacacaaaaatcgttaattgcctctaaatgtaagcaccatgactctagtctcttcttcttct 1620
Db 1561 AACACAAAATACGTAATGCCTCTTAATGTAAGCACCACCATGACTAGTGTCTTTGGTGTAT 1620
Qy 1621 attcttacaacotttagtgaaaaatccttatttttcaatcatttccctccaagtggtaaca 1680
Db 1621 ATTCTTACATPACCTTTAGTGGAAAAATCCTTATTTTTCATCATGCGCCTCCCAAGTGGTAC 1680
Qy 1681 gtctgaagtctctctgtattacctgagaagcaagatggaactgttttctctctctctc 1740
Db 1681 GTTCTAAGATGCTCTCTTGTATTAACCTGAGAAAGCAAGATGGAACGTGTTTTCTCTCTC 1740
Qy 1741 tgttgcaaacactctctacatcctctctctcttattacaaattcacatgattctcttctac 1800
Db 1741 TGTTGCCAACATCCTCTACATCCTCTCTCTTCTTATTACAATTCACATAGTCTTTTATCAC 1800
Qy 1801 tgaattctcaacccaatatttctctcatttgaaccaggagtcagctgttactagcaactg 1860
Db 1801 TGAATTTCTCAACCAATATTCTCTACTATTGAACAGGAGTCAAGTGTACTAGCAACTG 1860
Qy 1861 ctccaaggaatagatgaaatccccccctccactctctctctctctctctcagcactgattta 1920
Db 1861 CTCCAAGGATAGATGATGAATAATCCCCCTCCACTTCTCTGTACGGACACCTCGAATCATTT 1920
Qy 1921 ttgtgtgtgaggagctggagaattctccacaaatgttcccaaatccttctctcagctg 1980
Db 1921 TTGTGTGTGAGGAAGCTGGAGAAATTCACCAAAATGTTCCCAAAATCCTTATCTCTCAGCT 1980
Qy 1981 tgaaggtaaaaattggaacatcactggaattgggttggaaactctgaaacaaagaaatttg 2040
Db 1981 TGAAGGTAAAAATTGGAACATCACTGGAATGGGGTGGAAACATCTGGAACCAAGAAATTTG 2040

Qy	2041	atgactctgtatacttagacacaagaagagtgtaaaactccgaagtctctaaatcag	2097
Db	2041	ATGACTCTGTGATACTTAGACCAAGCAGAGTGTAATAACTCCGAAGTCTCTAAATCAG	2097
RESULT ⁴	4		
AF001846	AF001846	3058 bp mRNA linear PRI 01-OCT-2001	
LOCUS		Homo sapiens lymphoid phosphatase Lyp1 mRNA, complete cds.	
DEFINITION			
ACCESSION	AF001846		
VERSION	AF001846.1	GI:4100631	
KEYWORDS		human.	
SOURCE			
ORGANISM		Homo sapiens	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE		1 (bases 1 to 3058)	
AUTHORS		Cohen,S., Dadi,H., Shaoul,E., Sharfe,N. and Roifman,C.M.	
TITLE		Cloning and characterization of a lymphoid-specific, inducible	
		human protein tyrosine phosphatase, Lyp	
JOURNAL		Blood 93 (6), 2013-2024 (1999)	
MEDLINE		99168989	
PUBMED		10068674	
REFERENCE		2 (bases 1 to 3058)	
AUTHORS		Roifman,C.M.	
TITLE		Direct Submission	
JOURNAL		Submitted (30-APR-1997) Immunology and Allergy, The Hospital For	
		Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8,	
		Canada	
FEATURES		Location/Qualifiers	
source		1..3058	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/cell_type="thymocytes"	
		42..2468	
CDS		/codon_start=1	
		/product="lymphoid phosphatase Lyp1"	
		/protein_id="AAD00904.1"	
		/db_xref="GI:4100632"	
		/translation="MDREILQKFLDEAOSKKIKKEFANEFLKLRQSKYKADKY	
		PTTVAENAKNKKRKYDILPYDSRVELSILTSDESSYINANFIKVGPKRAYIAT	
		QGPLSTLLDFWRMWEYSVLIIYMACWEKKEEMGRKCYEAWPEGMOLFPGPVSVC	
		BAERKSDYIIRTLKVFNSETRTIYQFHYKNWDHPDVPSSIDILFLINDVRCYQED	
		DSVPICHCSCAGRTGVCIAVDITWMLKDGIPENFSVFLIREMRTQPSLVOT	
		QEOVELYNAVLELFKQMDVIRDKHSGESQAKHCIPKRNHTLQADSYSPNLKSVIT	
		KAAKMNQRTKPIKEIKSSPDFRTSEISAKELVLHPAKSSTSFDELELNYSPDKNA	
		DTTWKQTKAPEIVGELOKHQSLDLSLLEPGCSNKPYNAAKRYENSKYPIRTKS	
		TPFELIQREKEVDSKENSFLESQPHDSCFVEMQAQVNHVSAELINTSLPIDSXH	
		QIRNASNVKHHDDSALGVYSIPLVENPYFSSWPPSGTSSKMSLDLPEKQDGVFSS	
		LLPTSSLSFYNSHDSLSNPTNISLLNQESAVLATAPRIDDIPLPVPRTPE	
		SFIVVEEAGEFSPNPKSLSSAVKVKIGTSLWEGTSEPKKDDSDVILRPSKVLR	
		PKSELHODRSSPPPLPERTLESFLEADDCMQASITETYSTYPTDMENSTSSKQFL	
		KTPCKSFTRSKLILRNMKKSICNSCPPNKPESVQSNSSSFLNFGFANRFSKPKG	
		PRNPPPTWNI*	
BASE COUNT	1015 a	592 c	553 g 898 t
ORIGIN			
Query Match	88.9%;	Score 2093.8;	DB 9; Length 3058;
Best Local Similarity	99.9%;	Pred. No. 0;	
Matches 2095;	Conservative 0;	Mismatches 2;	Indels 0; Gaps 0;
Qy	1	tccctcaacctactatagactattttctgtctgcagatgcacccaagagaaattc	60
Db	1	TCCTCAACCTACTTATAGACTATTTTCTGCTGTGAGCATGGACCAAGAGAAATC	60
Qy	61	tgcagaagtcttgatgagcccaagaagaaattactaaagagagtttgccaatg	120
Db	61	TGCAGAAGTCTCTGGATGAGGCCCAAGCAAGAAATTAATAAGAGAGGATTG	120
Qy	121	aatttctgaagtgaaggcaatctaccagaagagcagacaaaacctatctacaa	180
Db	121	AATTCTTGAAGCTGAAAGGCAATCTTACCAGATACAGGACAGACAAAACCT	180

Qy	181	ctgtggctgagaatgccaaagaatatcaagaaaaacagataataagatatatttgcctatg	240
Db	181	CTGTGGCTGCAATGCCAAGAATATCAAGAAAAACAGATATAAGATATTTCCTATG	240
Qy	241	attatagccgggtagaactatccctgataaacctctgtagagattccagctacatcaatg	300
Db	241	ATTATAGCCGGGTAGAACTATCCCTGTATAAACCCTGTGATGAGATTTCACAGTACAT	300
Qy	301	ccaacttcattaaaggagtttatggaccacaaggcttatattgccaccagaggtccttat	360
Db	301	CCAACTTCATTAAAGGAGTTTATGGACCCCAAGGCTTATTATGGCCACCAGGGTCT	360
Qy	361	ctacaacctctggactctgtggagatgatttgggaatatagtgctcttcatcttgta	420
Db	361	CTACAACCTCTCTGGACTCTCTGGAGGATGATTGGGAATATAGTCTTATCATTTGTA	420
Qy	421	tggcatgcatgagttatggaatggaagaaagtgagcgctactggctgagccaag	480
Db	421	TGGCATGCTGAGTATGAAATGGGAAAGAAAAAGTGTGAGCGCTACTTGGCTGAGCC	480
Qy	481	gagatgcagctggaatttggcccttctctgtatctgtgaagctgaagctgaagaaat	540
Db	481	GAGATGCAGCTGGAAATTTGGCCCTTCTCTGTATCTCTGAAGCTGAAAAAGGAAT	540
Qy	541	ctgattataatcaggaactctaaaagttaagtccaatagtgaaactcgaactatctacc	600
Db	541	CTGATTATATTAATCAGGACTCTAAAAGTTAAGTTCAATAGTGAAGTGAAGTATCT	600
Qy	601	agttcattacaagaattggccagaccatgatgtacctctcatatagacacctatcttg	660
Db	601	AGTTTCATTACAAGAAATGGCCAGACCATGATGTACCTTCATCTATAGACCTATTCTG	660
Qy	661	agctcatctggatgtaoigtgttaccagaagatgaagtggtcccatgcatctcaact	720
Db	661	AGCTCATCTGGGATGACGTTGTTACCAAGAGGATGACAGTGTCCCATATGCAATCT	720
Qy	721	gcagtgcgtgctggaaggactggttattgtgctattgttattacatgatgagt	780
Db	721	GCAGTGTCTGGCTGTGGAAGGACTGGTGTATTGTTGCTATTGTTGATTATACATG	780
Qy	781	tgctaaagatgaggaataatctctgagaacttcagtggtttcagtttgatccgggaaatgc	840
Db	781	TGCTAAAAAGATGGGATAATTCCTGAGAACTTCAGTGTTCAGTTTGTATCGGGAAT	840
Qy	841	ggacacagagccttcattagttcaaacgcaggaacaatatgaactggtctacaaatgctg	900
Db	841	GGACACAGAGCCTTCATTAGTTTCAACCGCAGGAACATATGAAGTGGTCTACATGCTG	900
Qy	901	tattagaactatttaagagacagatggatgttatcagagataaaacattctctggaacagag	960
Db	901	TATTAGAACTATTTAAGACAGATGGATGTTATCAGAGATAAACATCTCTGGAACAGA	960
Qy	961	gtcaagcaaaagcattgtattcctgagaaaaatcaacactctccaagcagactcttcttc	1020
Db	961	GTCAAGCAAAAGCATTGTATTCTGTAGAAAAATCACACTCTCCAAGCAGACTCTTTAT	1020
Qy	1021	ctaatctaccaaaaagtagcaaaaagcagcaaaaatgatgaacaaacaaagacaaaaa	1080
Db	1021	CTAAATTTACAAAAAGTACCACAAAAGGACGCAAAAAATGATGACCCACAAAGGACA	1080
Qy	1081	tggaatacaagaatcttctcctttgacttttaggactctgaaataaagtgcaaaagaag	1140
Db	1081	TGGAATACAAGAAATCTTCTCTTCTTACTTTAGGACTTCTGAAATAAGTGCAAAAAG	1140
Qy	1141	agctagttttgcacccctgtaaaatcaagcactctcttgcactttctggagctaaattaca	1200
Db	1141	AGCTAGTTTTCACCCCTGTAATCAAGCATTCTTTTGTACTTTCTGGAGCTTAATATCA	1200
Qy	1201	gttttgacaaaaatgctgacacacatgaatggaagcagacaaagcatttccaatagttg	1260
Db	1201	GTTTTCACAAAAATGCTGACACACATGAATGCGACAGCAAGGCATTTCCAAATAGTTG	1260

Qy 1261 gggagcctcttcagagcatcaaaagtgtggattgggctctcttttggtaggaggtt 1320
Db 1261 GGGAGCCTCTTCAGAAAGCATCAAAAGTTTGGATTGGGCTCTCTTTTGGAGGATGT 1320
Qy 1321 ctaattctaaacctgttaaatgcagcaggaagaatatattttaaattcaaaaggcgcaataaacac 1380
Db 1321 CTAATTCCTAAACCTGTAAATGCAGCAGGAAGATATTTTAAATCAAAAGTGCCTAAATACAC 1380
Qy 1381 ggaccacaactcctcttgaattgatcacgagagagaaccaaaggaggtgacacga 1440
Db 1381 GGACCAATCAACTCTCTTTGAAATGATACAGCAGAGAGAAACCAAGAGGTGGACACGA 1440
Qy 1441 aggaacactttcttatttggaaatctcaaccacatgattctgttttttagagatgcagg 1500
Db 1441 AGGAAACCTTCTTATTTGGAATCTCAACCACATGATTCCTGTTTGTAGAGATGCAGG 1500
Qy 1501 ctcaaaaagtgaatgctgtttcttcagcagaactgaattatctcaactgcacatgactcta 1560
Db 1501 CTCAAAAAGTAAATGCATGTTCTTTCAGCAGAACTGAATTAATTCATGCTGCATATGACTCTA 1560
Qy 1561 aacccaaatagcgttaatgcctctaatgtaaagcaccatgactctagtgctcttggtgat 1620
Db 1561 AACACCAATAGTAATGCCCTCTAATGTAAAGCACCATGACTCTAGTGTCTCTTGGTGTAT 1620
Qy 1621 attctacataccttttagtggaatccttatttttcaatcatgacctccaaagtgtgacca 1680
Db 1621 ATTCTTACATACCTTTAGTGGAAATCCTTATTTTTCATCATGCTGCCTCAAAGTGGTACCA 1680
Qy 1681 gttctaaagatctctgtatttaccttgagaagaagaatggaactgttttttctctctctc 1740
Db 1681 GTTCTAAGATGTCTTGTATTACCTGAGAAAGCAAGATGGAACGTGTTTTCCTCTCTCTC 1740
Qy 1741 ttttgccaaactcctctacatccctctctctcttatttacaattcacatagttctttatcac 1800
Db 1741 TGTGTCACATCTCTTACATCCCTCTCTCTTATTACAAATACATGATCTTTATCAC 1800
Qy 1801 tgaattctccaaaccaatatttctcactatgtgaaccaggagtcagctgtactagcaactg 1860
Db 1801 TGAATTCCTCAACCAATATTTCTCTCACATTTTGAACCAAGAGTCACTGTACTAGCAACTG 1860
Qy 1861 ctcaagagatagatgaataatccccctccactctcactctctgtacgcagacactgaactta 1920
Db 1861 CTCCAAGGATAGATGAATGAAATCCCCCTCCACTTCTGTACGACACACTGAAATCATTTA 1920
Qy 1921 ttgtgttgaggaagctggagaattctcaccaaatgttcccaaatcccttactcactcagctg 1980
Db 1921 TTGTGTTGAGAGCTGGAGAAATCTCACCAAATGTTCCCAATCCCTATCCCTCAGCTG 1980
Qy 1981 tgaagggtaaaaattggaacatcactggaatggggtggaacatctgaaccaaagaatttg 2040
Db 1981 TGAAGGTAAAAATTGGAACATCACTGGAATGGGTGGGAACATCTGAACCAAGAAATTTG 2040
Qy 2041 atgactctgtatacttagaccagaagagtgtaaaactccgaagtcctaataacag 2097
Db 2041 ATGACTCTGTGATCTTAGACCAAGCAAGAGTGTAAAACTCCGAAGTCTTAAATCAG 2097

RESULT 5
AR14916 AR14916 3580 bp DNA linear PAT 08-AUG-2001
LOCUS AR14916 Sequence 1 from patent US 6228641.
DEFINITION AR14916
ACCESSION AR14916
VERSION AR14916.1 GI:15114507
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 3580)
AUTHORS Jallal, B. and Plowman, G.D.
TITLE Diagnosis and treatment of PTP04 related disorders
JOURNAL Patent: US 6228641-A 1 08-MAY-2001;
FEATURES Location/Qualifiers
source 1. 3580

BASE COUNT 1213 a 666 c 618 g 1083 t
ORIGIN
Query Match 88.1%; Score 2074.6; DB 6; Length 3580;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2090; Conservative 0; Mismatches 4; Indels 3; Gaps 1;
Qy 1 tccctcaaccttaactatagactatttttctgtctgcagcatggaccacaagaataatc 60
Db 12 TCCCTCAACCTACTATAGACTATTTTCTTCTCTGCATGACCAAAAGAGAAATTC 71
Qy 61 tgcagaagtctctgagtgagcccaagcaagaaataactaaagagaggtttgccaatg 120
Db 72 TGCAGAGTTCCTCGATGAGGCCCAAGCAAGAAATTAATTAAGAGAGGAGTTTCCCAATG 131
Qy 121 aattctgaagctgaaggccaatctaccaaagtacaaggcagacaaacctatctctcaa 180
Db 132 AATTTCTGAAGCTGAAAAGGCAATCTACCAAGTACAAGGCAGACAAACCTATCTACAA 191
Qy 181 ctgtggtcgagaatgccagaataatcaagaaaaacagataataaggaatatatttgcctatg 240
Db 192 CTGTGGCTGAGAACCCCAAGAAATATCAAGAAAAACAGATATAAGGATATTTTGGCCCTATG 251
Qy 241 attatagccgggtagaactatccctgataacctctgtagagattccagctcacatcaatg 300
Db 252 ATTATAGCCGGGTAGACTATCCCTGATACCTCTGTATGAGATTCAGCTACATCAATG 311
Qy 301 ccaacttcattaaaggaggtttatggccccagggttatatggccccagggtcccttat 360
Db 312 CCAACTTCATTAAGGGAGTTTATGGACCCCAAGGCTTATATGGCCACCAGGCTCTTTAT 371
Qy 361 ctcaacccctctgactctgagagattgttgggaatatagttccttatacttatttta 420
Db 372 CTACAAACCTCCTGGACTTCGTGGAGGATGATTTGGGAATATAGTGTCTTATCATATTGTTA 431
Qy 421 tggcatgcagtagatgaatggaaagaaagtgctgagcgtactgggctgagccag 480
Db 432 TGGCATGCATGAGATGAATGGAAAGAAAGAGTGTGAGCGCTACTGGCTGAGCCAG 491
Qy 481 gagagatgcagctggaatttggcccttctctgtatcctgtgaagctgaaaaaggaaat 540
Db 492 GAGAGATGCAGCTGGAATTTGGCCCTTCTCTGTATCTGTGAAAGCTCAAAAAAGGAAAT 551
Qy 541 ctgattataaatcagagactctaaaagttaagtccaatagtgaaactcgaaactatctacc 600
Db 552 CTGATTATATAATCAGGACTCTAAAAAGTTAAGTTCAATAGTGAAGCTCGAAACTATCTACC 611
Qy 601 agtttcattacaagaattggccagaccatgatgtacctctcatctatagaccctattctg 660
Db 612 AGTTTCATTACAAGAAATGGCCAGACCATGATGTACCTTCATCTATAGACCTTATTTCTTG 671
Qy 661 agtcatctgggagtacgtttgtaccaagagagatgacagtggttcccatatgcattcact 720
Db 672 AGCTCATCTGGGATGTACGTTGTTTACCAAGAGGATGACAGTGTTCCTCATATGCATTCAC 731
Qy 721 gcagtgcgtgtggaaggactggttatttctgtctattgttattgtattgattatatacatgagt 780
Db 732 GCATGTGTGGCTGTGGAAGGACTGGTGTATTGTGTGTA---TTGATATATACATGGGATGT 788
Qy 781 tgcataaagatgggaataattctctgagaacttcagtttttcagtttgcagtcggggaatgc 840
Db 789 TGCTAAAGATGGGATTAATCTCTGAGAACTTCAGTGTGTTTTCAGTTCGCGGAATATGC 848
Qy 841 ggacacagagccttctattagttcaaacgcaggaggaacaatatgaactgggtctacaatgctg 900
Db 849 GGACACAGAGGCCCTTCATTAAGTTCAGGAGGAAACAAATATGAATCGTCTTACAAATGCTG 908
Qy 901 tattagaactatttaagagacagatgattgtttatcagagataaacattcttgggaacagaga 960
Db 909 TATTAGAACTATTTAAGAGACAGATGGATGTTATCAGAGATAAACAATCTTGGAGACAGAGA 968

QY 961 gtaagcaagcattgtattcttgagaaaaatcacactctcaagcagactcttattctc 1020
Db 969 GTCAAGCAAGCATTGTATTCTCTGAGAAAAATCACACTCTCAAGCAGACTCTTATTCTC 1028
QY 1021 ctaatttaccaaaagtaccacaaaaagcagcaaaaaatgatgaaccacaaaagacaaaa 1080
Db 1029 CTAATTTACCAAAAGTACCACAAAAGCAGCAAAAATGATGAACCAACAAAGGACAAAAA 1088
QY 1081 tggaaaataaagaattctctcttgacttttaggactctgaataaagtcaaaaagaag 1140
Db 1089 TGGAAATCAAGAAATCTCTCTTTGACTTTTAGGACTTCTGAATAAGTCAAAAGAG 1148
QY 1141 agtagttttgcacctgtcaataacagcactctcttttgactttctggagctaaaattaca 1200
Db 1149 AGCTAGTTTGGACCTGCTAATAATCAAGCACTTCTTTTGACTTTCTGGAGCTAAATTACA 1208
QY 1201 gttttgacaaaaatgctgacacaaacatgaaatggcagacaaaaggcatttccaatagtgg 1260
Db 1209 GTTTTGACAAAAATGCTGACACAAACCATGAATGGCAGACAAAGGCATTTCCAATAGTTG 1268
QY 1261 gggagcctctcagaagcattcaaaagtgttgatttggctctcttttggtagggatgtt 1320
Db 1269 GGGAGCCTCTCAGAGCAATCAAAAGTTTGATTGGGCTCTCTTTTGTGGGGATGTT 1328
QY 1321 ctaattctaaacctgtaaatgcagcaggaagatatttttaattcaaaaggtgccaaatacac 1380
Db 1329 CTAATTTCAACCTGTAATGTCAGCAGGAAGATATTTTAATTCAAAGGTGCCAATAACAC 1388
QY 1381 ggacaaaatacaactccttttgattgatcacagcagagagaaacaaaggagggtgacagca 1440
Db 1389 GGACCAAAATCAACTCTCTTTTGAATTGATACGACAGAGAAACCAAGAGGTTGGACAGCA 1448
QY 1441 aggaacactttctatttgggaatctcaaccatgatcttctgtttttagagatgcagg 1500
Db 1449 AGGAAACTTTCTTATTGGAAATCTCAACCACATGATCTCTGTTTGTAGATGCAGG 1508
QY 1501 ctcaaaaagtaatgcattgtttcttcagcagaactgaattattcaactgccatatgactcta 1560
Db 1509 CTCAAAAGTAATGTCATGTTTCTTCAGCAACTGAATATTTCACATGACATGACTCTA 1568
QY 1561 aacacaaaatagtaatgcctctaataatgaaagcacacatgactctagctctgttgggtgat 1620
Db 1569 AACACCAAAATACGTAATGCTCTTAATGTAAAGCACCATGACTCTAGTGTCTCTTGGTGTAT 1628
QY 1621 attcttacaacactttagtggaaaatccttatttttcatcattgctcctcaagtggtacca 1680
Db 1629 ATTTCTACATACCTTTAGTGGAAATCCTTATTTTTCATCATGCGCTCCAAAGTGTACCA 1688
QY 1681 gttctaagatgtctcttgattttacctgagaagcaagatggaactgttttctctctctc 1740
Db 1689 GTTCTAAGATGTCCTTGATTTACCTGAGAAGCAAGATGGAAGTGTTTTCTCTCTCTC 1748
QY 1741 tgttgcaaacactctctacatacctctctctctctattacaaattcacatagttcttttctac 1800
Db 1749 TGTGTGCAACATCCTCTACATCCTCTCTCTTCTATTATACAAATTCACATGATCTTTATCAC 1808
QY 1801 tgaattctcaacccaattctctcactattgaaccagagtcagctgtactagcaactg 1860
Db 1809 TGAATTTCCACCAATATTTTCTCTACTATTGAACAGAGTCAGCTGTACTAGCAACTG 1868
QY 1861 ctccaagtagatgatgaaatccccctccactctctctgaagcacactgaaatcatatta 1920
Db 1869 CTCCAAGGATAGATGATGAATCCCCCTCCACTTCTCTGTACGACACACCTGAAATCATTTA 1928
QY 1921 ttgtggttgaagagctggagaattctcaccaaaatgttccccaaatccttctcctcagctg 1980
Db 1929 TTTGTGTTGAGGAAGCTGGAGAAATTTCTACCAAAATGTTCCCAATCTTTATCTCAGCTG 1988
QY 1981 tgaaggttaaaaattgaaacatcatcgaatgggttgaacatctgaacaaaagaatttg 2040
Db 1989 TGAAGGTAAAAATTTGAAACATCACTGGAATGGGGTGGAACTCTGAACCAAGAAATTTG 2048
QY 2041 atgactctgtgacttagaccagaagcagtgtaaaactcgaagtccttaataatcag 2097

Db 2049 ATGACTCTGTGATACTAGACCAAGCAGAGTGTAACAACTCCGAAGTCTTAATCAG 2105
RESULT 6
AF150732 2877 bp mRNA linear PRI 04-MAY-2000
LOCUS Homo sapiens protein tyrosine phosphatase (PTP) mRNA, complete cds.
DEFINITION
ACCESSION AF150732
VERSION AF150732.1 GI:7688662
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2877)
AUTHORS Huang,C., Zhang,C., Wu,T., Peng,Y., Gu,Y., Zhang,L., Jiang,C.,
Li,Y., Han,Z., Wang,Y., Chen,Z. and Fu,G.
TITLE A novel gene expressed in human adrenal gland
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2877)
AUTHORS Huang,C., Zhang,C., Wu,T., Peng,Y., Gu,Y., Zhang,L., Jiang,C.,
Li,Y., Han,Z., Wang,Y., Chen,Z. and Fu,G.
DIRECT SUBMISSION
TITLE Submitted (12-MAY-1999) Chinese National Human Genome Center at
JOURNAL Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, P. R. China
FEATURES
source
1..2877
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="adrenal gland"
gene
1..2877
/gene="PTP"
CDS
71..2470
/gene="PTP"
/codon_start=1
/product="protein tyrosine phosphatase"
/protein_id="AAF67472.1"
/db_xref="GI:7688663"
/translation="MDQREILQKFLDEAQSKKITKEEFANERFLKLRKSTKYKADKTY
PTTVAEPKNINKNRYKIDPYDSRVELSLTSDSSYINANIKGYGPEAKIAT
QGPSTLLDFWRMTWEYSVLINMACMEYEMGKKCRKYAEPMGEMOLEGFPFVSVC
EAEKRKSDYIIRTLKFNSETRTIYQFHNWPDHVPSSIDPDLIELIWDVRCQTQD
DSVPTCIHCSAGCGRTGICADYTWMLIKGCIIPENFSVSLIREMTQRSLVQTQ
EYELVYNVLELPKROMDYIRDKHSGTSEQAKHCIPKNIHTLQADSYSPNLPKSTTK
RAKMNQORTKWEIKESSDFPRTSEISAKEDVSLAPENQAFPCISEVNSSPAKMLPH
HMPAKAFQIVGSSEASKLNWASLCIGNSNLVCSKIFNPQRCYPDQINSFELIQ
QENQGDGSKTFLFESQPHDSCFVEMQAKVNHVSSALNYSPLDYDSKHQIRNASNVK
HDDSSALGVYSYIPLVNPYFSSWPPSGTSSKMSLDLPKODGTVPFSSLLTSTSL
FSYYNHSDLSLNSPTNISRTQESAVLATAPRIDDEIPLHFLYGTPESTFVVEEAG
FESPVPKSLSSAVKVKICTSLWGGTSEPKKFDSDVILLRPSKSVKLRSKPELHQDR
SSPPPLPERLESFLADEDCMQAQSIETYSTSYPTDMENSTSKQTLKTPGKSFTR
SKSLKILRNMKKSCINCSPPNKAESVSNSSSFLNFGFANRFSPKPRPPPTWN
I"
BASE COUNT 944 a 583 c 523 g 827 t
ORIGIN
Query Match 74.7%; Score 1761; DB 9; Length 2877;
Best Local Similarity 94.0%; Pred. No. 0;
Matches 1978; Conservative 0; Mismatches 100; Indels 27; Gaps 13;
QY 1 tccctcaactctatagactattttcttctgcagctggaccacaaagagaatic 60
Db 30 TCCTCAACCTACTTATAGACTATTTTCTGCTCTGCAGCATGGACCAAGAGAATTC 89
QY 61 tgcagaagttcttgatgagggcccaagcaagaaaaattactaaagagagatttgc 120
Db 90 TGCAGAAGTTCCTGGATGAGGCCCAAGCAAGAAAATTACTAAAGAGGAGTTTGC 149
QY 121 aattcttgaagctgaaaaaggaactctacaagtcacaaggcagacaaaacctct 180

Db 150 AATTTCTGAAGCTGAAAGGCAATCTACCAAGTACAAAGCAGACAAAACCTATCTCTACAA 209
 Qy 181 ctgtgctgagaaatgccaaagaatatacaagaaacacagataaagatatatttgccctatg 240
 Db 210 CTGTGGCTGAGAGCCCAAGAAATATCAGAAACACAGATATAGGATATTTTGCCTATG 269
 Qy 241 attatagccgggtagaactatccctgataaactctgatgaggtattccagctacatcaatg 300
 Db 270 ATTATAGCCGGGTAGAACTATCCCTGTAGTAACCTCTGTATGAGGATGCCAGCTTACATCAATG 329
 Qy 301 ccaacttcaataagggagatttataggacccaaaggcttatattgcaaccacagggctctttat 360
 Db 330 CCAACTTCATTAAGGGAGTTTATGGACCCGAGGCTTATATTGCCACCCAGGCTCCTTTAT 389
 Qy 361 ctacaacccctcctgacctctggaggatgatttgggaatatagtgctcttatcatgttta 420
 Db 390 CTACAAACCTCTCGACTTCTGGAGGATGATTTGGGAATATAGTGCTCTTATCATATGTTA 449
 Qy 421 tggcatgcatggagatgaaatgggaagaaagaaagtgtgagcgtactgagggctgagccag 480
 Db 450 TGGCATGCTGAGTATGAATGGAAAGAAAGAAAGTGTGAGCGCTACTGGGCTGAGCCAG 509
 Qy 481 gagagatcagctggaatttggcccttctctgtatcctgtgagcgtgaaagaaagaaat 540
 Db 510 GAGAGATCAGCTGGAATTTGGCCCTTTCTCTGTATCCTGTGAAGCTGAAAAAAGGAAT 569
 Qy 541 ctgattataatcaggactctaaaagttaagttcaatagtgaaactcgaaactatctacc 600
 Db 570 CTGATTATATATCAGGACTCTAAAGTTAAGTTCAATAGTGAAGCTGAAACTATCTATACC 629
 Qy 601 agtttcattacaagaattggccagaccatgatgtacccttcctctatagaccctattcttg 660
 Db 630 AGTTTCAATACAAGAAATGGCCAGACCATGATGTACCTTCATCTATAGACCCCTATTCTTG 689
 Qy 661 agctcatcgggatgtaagttgttaccgaagagatgacagtggttcccatatgcattcaact 720
 Db 690 AGCTCATCTGGGATGTCGGTTGTACCCAAAGAGGATGACAGTGTTCCTCCATATGCAATCACT 749
 Qy 721 gcagtgcgtgctggaagactggttatttctgtctattgtctattgttgattatcatggatg 780
 Db 750 GCAGTGCCTGGCTGTGGAAGGACTGGTGTATTGTGCTA---TTGATTATACATGGATGT 806
 Qy 781 tgcataaagatgggataattcctgagaacttcagtggttttcagtttgatccgggaaatgc 840
 Db 807 TGCTAAAGATGGGATATCTTCGAGACTTCAGTGTTCAGTGTGATCGGGAAATGC 866
 Qy 841 ggacagagagccttcattagttcaaacgcgaggaacaatatgaaactggtctacaatgctg 900
 Db 867 GGACACAGAGGCTTCATTAGTTCAAAACGACGAAACAATATGAAGTGTCTTACAATGCTG 926
 Qy 901 tattagaactattaagagacagatggatgttatcagagataaacattcttggaacagaga 960
 Db 927 TATTAGAACTATTAAAGACAGAGATGGATGTTATCAGAGATAAACAATTCGGAACAGAGA 986
 Qy 961 gtcaagcaaacgattgtattcctgagaaatacacactctccagcagacactcttattctc 1020
 Db 987 CTCRAAGCAACGATGTATTCTTGAGAAATATCACACTCTTCAGCAGACTCTTATCTC 1046
 Qy 1021 ctaatttaccataaagatgaccacaaagcagcaaaaatgatgaaaccaaaagggacaaaa 1080
 Db 1047 CTAATTTACCAAAAAGTACCACAAAAGCAGCAAAAATGATGAACCAACAAAAGGACAAAA 1106
 Qy 1081 tggaaatcaagaactcttctccttgacttttagactcttgaaatagtgcaaaagaag 1140
 Db 1107 TGGAAATCAAGAAATCTTCTCTTGTGACTTTTAGACTTTCGAATAAGTCAAAAGAAAG 1166
 Qy 1141 agctagttttgaccctgctcaaatcaagcactctcttttttgactttcttgagcgtataattaca 1200
 Db 1167 ACGTAAGTTTGGCCCGG---AAACACAGGCTTTTGTTCCTTCGGAAGTAATTTCCA 1223
 Qy 1201 gttttgacaaaaatgctgacacacacatgaaatggcgacagaagggcatttccaatagttg 1260
 Db 1224 GTTTTGT---CAAAAATGTTGCCCCACCATGAATGCCAG-CAAAGGCTTTCCAAATAGT-- 1278

Qy 1261 gggagcctcttcagaagcatcaaaagtttggatttgggctctcttctgttggagggatgt 1320
 Db 1279 GGGAGACTCTTCAGAGCATCAAG-TTGAATTTGGGCTCTCTTTGTTT---AGGAATGT 1334
 Qy 1321 ctaattctaaacctgtaaatgcagcaggaagatatatttaattcaaaagtgcaataaacac 1380
 Db 1335 CTAATTTCTAACCTTGT--ATGCAGCAGAAAGATATTT--AATCCAAGGTGCCAAT-ACCC 1389
 Qy 1381 ggaccaaatacaactccttttgaattgatacagagagagagaaacaaagaggtggacagca 1440
 Db 1390 GGACCAATCAACT-CTTTTGAATTGATCCAGCA---AGAGAACAAGGAGGTGACAGCA 1445
 Qy 1441 agaaaaactttcttatttggaaatctcaaacacatgatcttctgtttagagatgcaag 1500
 Db 1446 AGAAAACCTTCTTATTT---GAATCTCAACCAACATGATCTTGTGTTGTAGAGATGAGG 1502
 Qy 1501 ctcaaaagtaatgcattgtttcttcagcagaaactgaattattcactgcacatgactctata 1560
 Db 1503 CTCAAAAGTAAATGATGTTTCTTCAGCAGAACTGAATTAATTCATGTCATATGACTCTA 1562
 Qy 1561 aacacaaaatcgttaatgcctcttaattgaaagacacatgactctagctcttctgttat 1620
 Db 1563 AACACCAAAATACGTAATGCCCTTAATGTAAAGCACCATGACTCTAGTGTCTCTTGGTGTAT 1622
 Qy 1621 attctacatacctttagtggaataccttatttttcatcatggtccctcgaagtgttacc 1680
 Db 1623 ATTCTTACATACCTTTAGTGGAAATCTTATTTTTCATCATGCTCCCAAGTGGTACCA 1682
 Qy 1681 gttctaagatgctcttctgatttacctgagaagcaagatggaaactgttttctctctctc 1740
 Db 1683 GTTCTAAGATCTCTCTTGTATTTACCTGAGAAGCAAGATGGAACGTGTTTCTCTTCTC 1742
 Qy 1741 tgttgcaaacatctctacatccctctctctctctctcttattacaaatcacatagttctttacac 1800
 Db 1743 TGTTGCCAACATCTCTACATCCCTCTTCTTATTATTAATTCATGATTCATTTATCAC 1802
 Qy 1801 tgaattctccaaaccaatatttctcactattgaaagagagctcagctgactagcaactg 1860
 Db 1803 TGAATTTCCAAACCAATATTTTCTCAGTATTGAACAGGAGTCTAGCTGTACTAGCAACTG 1862
 Qy 1861 ctcaagagatagatgataatccccctccactctctctgtacgcagacactgaatcattta 1920
 Db 1863 CTCAAGGATAGATGATGAATCCCTCCACTTCTCTGTACGGGACACCTGAATCATTTA 1922
 Qy 1921 ttgtggttgaggaagctggagaattctcacaagaatgttcccaaatcttcttccctcagctg 1980
 Db 1923 TTGTGGTTGAGGAAGCTGGAGATTTCTACCAAAATGTTCCCAAAATCTTATCTCCAGCTG 1982
 Qy 1981 tgaaggttaaaatttgaaacatcactggaatgggtggaacatctgaacaaagaaatttg 2040
 Db 1983 TGAAGGTAAAAATTTGGAAACATCACTGGAATGGGTGGAAACATCTGAAACCAAGAAATTTG 2042
 Qy 2041 atgactctgtatcttagaccaagcagagtgtaaaactccgagctctaaatcaggtta 2100
 Db 2043 ATGACTCTGTGATCTTAGACCAAGCAAGAGTGTAAAACCTCCGAAGTCTTAAATCAGAAC 2102
 Qy 2101 aaat 2105
 Db 2103 TACAT 2107

RESULT 7

MUSPROTYPH
 LOCUS
 DEFINITION
 Mouse protein tyrosine phosphatase (70zpep) mRNA, complete cds.
 ACCESSION
 M90388
 VERSION
 M90388.1 GI:200522
 KEYWORDS
 protein-tyrosine phosphatase.
 SOURCE
 Mus musculus
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

MUSPROTYPH 2734 bp mRNA linear ROD 27-APR-1993
 DEFINITION Mouse protein tyrosine phosphatase (70zpep) mRNA, complete cds.

Db 1569 GGGAGCTGCAGGTGCAAAAGAGTGGCCCATGTTCTTTCAGAAAGAGCTGAATTATTACATGCG 1628
Qy 1549 catatgactctaaacaccaaatacgtactcctaaatgtaaaagcaccatgactctagtq 1608
Db 1629 CTGGTGCTT-----GTGATGGCTGTGTGTGCGCCGGCACAGCCCGCGG 1673
Qy 1609 ctcttggtgtattcttacatcaccttttagtggaaatcccttattttcatcatcgccctc 1668
Db 1674 CTTTGAGAGTGTCATCTGTACACATCTTTAGCGGAAGATCCCTTATTTTTCATCATCCCTC 1733
Qy 1669 caagtggtaccagttctcaagtgtctctgtatttaccttgagaagaagaatggaactgttt 1728
Db 1734 CCAATAGTCTGCTGATTCGAAGATGCTTTTGTATCTGCTCTGAGAAACAGGATGAGCCACTT 1793
Qy 1729 ttccttctctctgttgccaaatcctctacatccctctctcttatttattacattcacata 1788
Db 1794 CCCCTGGGCTCTATTGGCAGCTCTTCTACACCTCTCTTTTATYAGCAACCCACACAG 1853
Qy 1789 gttcttttacctgaattctccaaacaaatatttctcactatttgaaccaggagtcagctg 1848
Db 1854 ACTCCCTAGTATGAACACTCTGACAGCTTTTCCCCACCCTTAAACCAAGAGACAGCTG 1913
Qy 1849 tactagaactgctccaaagtagatgatgaataccccccctccacttctctgtagcgacac 1908
Db 1914 TAGAAGCTCTTCTCGGAGGACAGATGATGAATCCCGCCGCCACTCCTCGAACGGACAC 1973
Qy 1909 ctgaatcatttattgtgtgaggaagctggagaattctcaccaaatgttcccaaatctc 1968
Db 1974 CCGAGTCTTTTATGTGTGTGAGGAAGCCGGAGAGCCCTCACCAGCTGTTACCGAATCCT 2033
Qy 1969 tatctcagctgtgaagtaaaattggaacatcactggaatgggtgggaacatctggaac 2028
Db 2034 T-----ACCTCTGTGTGAACATTTGGAGCATCACCAGATGCAGTGGGACATCTGA-- 2085
Qy 2029 caaagaatttgatgactctgtgacttagaccagaagaagtgtaaaactcogaagctc 2088
Db 2086 -AATGAAGAGCCATGACTCTGTAGGTTTACACCAAGCAAGAATGTGAACATCCGAAGTC 2144
Qy 2089 ctaaatcaggtaaaaat 2105
Db 2145 CCAATCAGATCGACAT 2161

RESULT 8
LOCUS AL137856/c
DEFINITION Human DNA sequence from clone RP5-107303 on chromosome 1p13.1-13.3,
complete sequence.
ACCESSION AL137856
VERSION AL137856.24 GI:14970654
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 143618)
Glithero.R.
Direct Submission
Submitted (17-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
REQUEST: clonerquest@sanger.ac.uk
On Jul 19, 2001 this sequence version replaced gi:14586004.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em., EMBL; Sw.,
SWISSPROT; Tr., TREMBL; Wp., WORMPEP; information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chrl
RP5-107303 is from the library RPCI-5 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm

VECTOR: pCYPAC2

This sequence is the entire insert of clone RP5-107303 The true
right end of clone RP11-215C20 is at 30467 in this sequence. The
true right end of clone RP11-324J2 is at 119169 in this sequence.

FEATURES

source

1. 143618
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="p13.1-13.3"
/clone="RP5-107303"
/clone_lib="RPCI-5"
257..322
/note="MER5A repeat: matches 41..109 of consensus"
763..1053
/note="Alusx repeat: matches 1..312 of consensus"
1348..1698
/note="L2 repeat: matches 2070..2417 of consensus"
1919..1969
/note="L2 repeat: matches 2655..2705 of consensus"
2305..2446
/note="MIR repeat: matches 40..176 of consensus"
5652..5927
/note="L1ME repeat: matches 5524..5813 of consensus"
6295..6428
/note="BC200 repeat: matches 2..147 of consensus"
6653..6700
/note="8 copies 6 mer acacac 97% conserved"
7925..7980
/note="28 copies 2 mer aa 75% conserved"
8071..8369
/note="Alusx repeat: matches 1..296 of consensus"
9048..9342
/note="Alusx repeat: matches 1..298 of consensus"
9793..10559
/note="L1MC1 repeat: matches 5526..6320 of consensus"
10574..10730
/note="MIR repeat: matches 95..262 of consensus"
10759..10907
/note="MIR repeat: matches 2..139 of consensus"
10908..11219
/note="AluJo repeat: matches 1..296 of consensus"
11220..11315
/note="MIR repeat: matches 139..237 of consensus"
12210..12464
/note="Alusg repeat: matches 1..291 of consensus"
12916..13149
/note="MER58B repeat: matches 70..341 of consensus"
13150..13435
/note="Alusx repeat: matches 1..298 of consensus"
13458..13805
/note="L1MD2 repeat: matches 4851..5213 of consensus"
13806..14115
/note="AluJb repeat: matches 1..299 of consensus"
14118..14430
/note="Alusx repeat: matches 1..312 of consensus"
14431..14563
/note="L1MD2 repeat: matches 5213..5314 of consensus"
14565..14859

repeat_region	/note="AluSq repeat: matches 1. .294 of consensus" 14861. .14888
repeat_region	/note="14 copies 2 mer ta 92% conserved" 14910. .15215
repeat_region	/note="LMD2 repeat: matches 5302. .5609 of consensus" 15230. .15560
repeat_region	/note="MER2 repeat: matches 2. .345 of consensus" 15361. .16146
repeat_region	/note="LMD2 repeat: matches 5643. .6210 of consensus" 16270. .16479
repeat_region	/note="MTR repeat: matches 48. .262 of consensus" 16840. .17032
repeat_region	/note="MER3 repeat: matches 2. .207 of consensus" 17472. .17501
repeat_region	/note="15 copies 2 mer ac 100% conserved" 18839. .19765
misc_feature	/note="Cpg island" /evidence=not_experimental
repeat_region	20385. .20649
repeat_region	/note="AluJo repeat: matches 1. .297 of consensus" 20642. .20734
repeat_region	/note="FAM_C repeat: matches 4. .116 of consensus" 22760. .23065
repeat_region	/note="AluSx repeat: matches 1. .306 of consensus" 23275. .23506
repeat_region	/note="AluJb repeat: matches 1. .225 of consensus" 23540. .23584
repeat_region	/note="L2 repeat: matches 2647. .2694 of consensus" 23590. .23671
repeat_region	/note="AluJ/FLAM repeat: matches 1. .83 of consensus" 23714. .24013
repeat_region	/note="AluSx repeat: matches 1. .297 of consensus" 24532. .24817
repeat_region	/note="AluJb repeat: matches 17. .293 of consensus" 25033. .25166
repeat_region	/note="FAM_C repeat: matches 11. .143 of consensus" 25261. .25549
repeat_region	/note="AluSc repeat: matches 1. .289 of consensus" 25675. .25877
repeat_region	/note="MLTJ repeat: matches 2. .214 of consensus" 25998. .26083
repeat_region	/note="Charliel repeat: matches 3. .92 of consensus" 27261. .27556
repeat_region	/note="AluSx repeat: matches 3. .289 of consensus" 28123. .28175
repeat_region	/note="L2 repeat: matches 2678. .2733 of consensus" 28395. .28708
repeat_region	/note="AluSg repeat: matches 1. .310 of consensus" 29346. .29766
repeat_region	/note="L2 repeat: matches 2085. .2448 of consensus" 29767. .30052
repeat_region	/note="AluSg repeat: matches 1. .287 of consensus" 30053. .30345
repeat_region	/note="L2 repeat: matches 2448. .2750 of consensus" 31191. .31486
repeat_region	/note="AluSg repeat: matches 1. .296 of consensus" 31804. .31932
repeat_region	/note="AluJb repeat: matches 1. .127 of consensus" 32054. .32075
repeat_region	/note="11 copies 2 mer aa 100% conserved" 32503. .32780
repeat_region	/note="AluJo repeat: matches 1. .288 of consensus" 32791. .32817
repeat_region	/note="9 copies 3 mer tgt 96% conserved" 32817. .33099
repeat_region	/note="AluSg repeat: matches 1. .282 of consensus" 33156. .33450
repeat_region	/note="AluJb repeat: matches 6. .301 of consensus" 33747. .34055
repeat_region	/note="AluRa repeat: matches 1.309 of consensus" 34584. .34775
repeat_region	/note="L1ME repeat: matches 5644. .5823 of consensus" 34786. .34930

	/note="FLAM_C repeat: matches 12. .142 of consensus" 34931. .35225 /note="LIME repeat: matches 5265. .5563 of consensus" 35257. .35727 /note="LIME repeat: matches 4762. .5233 of consensus" 35743. .36021 /note="ALUdJ repeat: matches 1. .287 of consensus" 36340. .36488 /note="L1PA15 repeat: matches 6009. .6157 of consensus" 37329. .37776 /note="L1MC5 repeat: matches 7275. .7680 of consensus" 38067. .38121 /note="L2 repeat: matches 2681. .2735 of consensus" 38346. .38564 /note="L1MB5 repeat: matches 5921. .6142 of consensus" 39036. .39268 /note="MIR repeat: matches 20. .262 of consensus" 39278. .39564 /note="AlusC repeat: matches 1. .294 of consensus" 39575. .39657 /note="ALUJo/FRAM repeat: matches 206. .289 of consensus" 41653. .41957 /note="ALUdJ repeat: matches 5. .293 of consensus" 42324. .42634 /note="AluY repeat: matches 1. .311 of consensus" 43028. .43056	Query Match 34.7%; Score 816.8; DB 9; Length 143618; Best Local Similarity 99.8%; Pred. No.le-177; Matches 818; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY	1035 agtaccacaagcagcaaaatgatgaaccaacaaaggacaaaatggaaatcaaagaa 1094	
DB	45496 AGTACCACAAGACGCAAAAATGATGAACCAACAAAGACAAAAATGSAATCAAAGAA 45437	
QY	1095 tcttttccttgacttaggacctctgtgaataaagtgcacaagaagactggttttgac 1154	
DB	45436 TCTTTCTCTTGACTTTAGGACTTCTGAAATAAGTCGAAAAGAAGAGACTAGTTTGAC 45377	
QY	1155 cctgtaataacagcactcttttgactcttgagcttaaattacagtttigacaaaaat 1214	
DB	45376 CCTGTAAATCAAGCACCTCTTTTTGACTTCTGGAGCTTAATATTCAGTTTTGCACAAAAT 45317	
QY	1215 gctgacaaaccatgaatggcagacaagcgattcccaatagttggggagccttcag 1274	
DB	45316 GCTGACACACCATGAATGGCAGACAAGGCATTCCAAATAGTTGGGGGCCCTTCAG 45257	
QY	1275 aagcatcaaatgttgattgggtctcttttgttgaggatggttctaattctaaact 1334	
DB	45256 AAGCATCAAAAGTTGGATTGGGCTCTCTTTTGTGTGGAGGATGTTCTAATTCTAAAGCT 45197	
QY	1335 gtaatcgagcaggaagatatatttaattcaaaagtgccaataaacacgagcaaatcaact 1394	
DB	45196 GTFAAATGCAGCAGGAAGATATTTTAATCAAAAGTGCCAATAAACACGACCAATCAACT 45137	
QY	1395 ccctttgaaigtacacagacagaaaccaaagggtggacagcaaggaaaaactttct 1454	
DB	45136 CCTTTTGAAATTGATACGACAGAGAACCAGAGAGGTGGACGACGACGAANAACHTTTCT 45077	
QY	1455 tatttggaaatccaaccacatgattcttgtttgttagagatcgaggtccaaaaagtaatg 1514	
DB	45076 TATTTGGAAATCTCAACCATGATCTTTGTTGTAGAGATCGAGGCTCAAAAAGTAATG 45017	
QY	1515 cagtttcttcagcaggaactgaattattcaactgccatatgactctaaacaccaaatacgt 1574	
DB	45016 CATGTTTCTTTCAGCAGACTGAATTTATTCTACCTGCCATATGACTCTFAAACACCAAAATAGT 44957	
QY	1575 aaagcctctaagtgaagaccacatgactctagtcctctgtgtgtatattcttacatcact 1634	
DB	44956 AATGGCTCTTAATGTAAGAGCACCATGACTAGTGTCTTGTGGTGTATATCTTTACAFACCT 44897	
QY	1635 tttagtgaaaatcccttattttcatcattggctccaagtggtaccagttctcaagatgtct 1694	

Db	44896	TTAGTGGAAATCCTTATTTTCATCATGGCCTCCAGTGGTACCAAGTCTTAAAGATGTCT	44837
Qy	1695	cttgatttacctgagagcaagatggagactgttttccctctctctctgttgccaaacac	1754
Db	44836	CTTGATTTACCTGAGAGCAAGATGGAACTGTTTTCTCTCTCTGTTGCCAACATCC	44777
Qy	1755	tctacatccctctctctctattacaattcacatagttcttttatcactgaaattctccaaac	1814
Db	44776	TCTACATCCCTCTCTCTTATTACAATTACATGATCTTTATCACTGAATTCCTCCAACC	44717
Qy	1815	aatatctccctcaactattgaaccaggagtcagctgtactag	1854
Db	44716	AATATTTTCCCTACATATTGAACCCAGGAGTCACTGTACTAG	44677

RESULT	9	
LOCUS	BC017785	
DEFINITION	Homo sapiens, clone MGC:22283 IMAGE:4666653, mRNA, complete cds.	
ACCESSION	BC017785	
VERSION	BC017785.1	
KEYWORDS	GI:17389491	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
AUTHORS	1 (bases 1 to 1817)	
TITLE	Strausberg, R.	
JOURNAL	Direct Submission	
REMARK	Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
COMMENT	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgabs-femail.nih.gov Tissue Procurement: DCTD/DTF cDNA Library Preparation: CLONTECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www-shgc.stanford.edu Contact: (Dickson, Mark) mcd@paxil.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.	

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 36 Row: j Column: 17
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA q1: 15619016.

```

FEATURES             source
Location/Qualifiers
1..1817
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:22283 IMAGE:4566653"
/tissue_type="Prostate, adenocarcinoma."
/clone_lib="NIH_MGC_60"
/lab_host="DH10B"
/notes="Vector: pDNR-LIB"
90..629
/codon_start=1
/product="Unknown (protein for MGC:22283)"
/protein_id="AAH17785.1"
/db_xref="GI:17389492"
/translation="MDREILQKFLDAQSKKTKKEEFANEFUKLRQSKTKYKADRTY
PTTVAEPKNIKNRYKIDLYDSVELSITSDSSYINANFTKGVYGPRAIYAT
QSLTLLDFWRMIWEYVSLIIIVMACMEYEMGKEAKRSDYIITLTKVFNVSIVI
LAQHTSLQNLFSQITPAHP"
660 a 292 c 286 g 579 t
BASE COUNT
ORIGIN

```

Query Match 19.0%; Score 447.2; DB 9; Length 1817;
 Best Local Similarity 99.3%; Pred. No. 1.6e-92;
 Matches 449; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy	1	tccctcaacctactatagacattttttctgtctcagcatgagccaaagaaaattc 60
Db	49	TCCCTCAACCTACITATAGACTATTTTCTTGCTCTGCAGCATGGACCAAGAAATTC 108
Qy	61	tgcagaagtctctgagtgagcccaagcaagaaaattactaaagagaggtttgccaatg 120
Db	109	TGCAGAAGTTCTCGATGAGSCCCAAAGCAAGAAATTTACTAAAGAGAGGTTTCCCAATG 168
Qy	121	aattctgaaagctgaaagggcaatctaccaagtcaaggcagacaaacctatctctacaa 180
Db	169	AATTCTGAAGCTGAAAGGCAATCTCAAGTAGTACAAGGCAGACAAAACCTATCTCTACA 228
Qy	181	ctgtggctgaaatgcacaaagaatcaagaaaaacagatatagaagatatatttggccctatg 240
Db	229	CTGTGGCTTGAGAGGCCCAAGAATATCAAGAAAAACAGATATAAGGATATTTTGGCCCTATG 288
Qy	241	attatagccgggttagaacatccctgataacctctgatgaagattccagctacatcaatg 300
Db	289	ATTATAGCCGGTAGACTATCCCTGATACCTCTGATGAGGATTCACGTACATCAATG 348
Qy	301	ccaaactcaatgaaggagtttatggacccaaggcttatattgccaccagggtccctttat 360
Db	349	CCAACTTCATTAAAGGAGTTTATGGACCAAGGCTTATATTGCCACCCAGGGTCTCTTAT 408
Qy	361	ctacaacctctcgactcttggaagtatttggaagatttggaagatatagtcctctateattgta 420
Db	409	CTACAACCTCTCGACTCTGAGGAGTAGATTTCGGGAATATAGTGTCTCTATCATCTGTTA 468
Qy	421	tggcatgcatgagtgagtaaatgggaagaaa 452
Db	469	TGGCATGCATGGAGTATGAATGGGAAGNA 500

RESULT 10
 HUMTPPEST 3160 bp mRNA linear PRI 16-JUN-1993
 LOCUS Human protein tyrosine phosphatase (PTP-PEST) mRNA, complete cds.
 DEFINITION
 accession M33425
 version M33425.1 GI:292408
 keywords protein tyrosine phosphatase.
 source Homo sapiens cDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 AUTHORS Yang,Q.C., Tonks,N.K. and Sommercorn,J.
 TITLE Cloning and expression of ptp-pest: a novel, nontransmembrane
 protein tyrosine phosphatase
 J. Biol. Chem. 268, 6622-6628 (1993)
 JOURNAL 93203262
 MEDLINE
 FEATURES Location/Qualifiers
 1. 3160
 /organism="Homo sapiens"
 /db_xref="taxon.9606"
 /cell_line="HeLa"
 30. .2372
 /codon_start=1
 /product="protein tyrosine phosphatase"
 /protein_id="AAA36529.1"
 /db_xref="GI:292409"
 /translation="MEQVEILRKFQIRVQAMKSPDHNGDNFARFMRLLRLTKYRT
 EKITYPATGKEENVKNKRYKDILPFDHSRKVLTKTPSQSDYINANIKIGVYKPKA
 YVATQGLPANTVIDFWMVWYVNIIVACREFEMKRCRWYSPFLSDYEDPTTFAP
 KISCDEQARTDYFIRLTLEFVNSRRLLYQFVVMVPHDVPSPSSYLDMLSLMRK
 YQBEDVPICIHCSAGCGRTGACIDYTNLLKAGKIPFENVFNLIQEMRTQRIISA
 VQTEQELVTHRAIQLEFKQLQYIEHQAQTDAGVNEINENIMSSIEPERQDSPP
 PKPPTRSCSLVEGDAKEEILQPEHPHPVILPTSPSPAFFVTVTVMQNDRYHPKPP
 LHMVSEQSLADKNRYSKTELPGLKNESTIEQIDKRLERNLSFEIKKVPQLQEGPKPS
 DGNLTNRGHAIKRISASPCIAQKISKPOELSDLNAGDTSQNSVDCSVQTSQNVSV

TPPEQSNDTPPPRDLPLDEKHVWTFHGPENAIPIPLSEGNSSDINYQTRKTV
SLTSPFTQVETPDLVDHNDTSPLRTPLSFTNPLHSDSDSDSDSDSGAVTQNTKNI
STASATVSAATSTESITKVLPMISIARHNIAGTTHSGAEKDVDSPPPLPERTP
ESFVLASHNTPVSEWSELOSEOEKSEGLITSENEKCDHPAGGIHYEMCIECP
PTFSKREQISENPTEADIGFNGRCRPGKPRDPSPSEWT"

BASE COUNT 1077 a 586 c 612 g 885 t
ORIGIN

Query Match		16.1%;	Score 378.2;	DB 9;	Length 3160;
Best Local Similarity		65.2%;	Pred. No. 1.2e-76;		
Matches 573;		Conservative 0;	Mismatches 303;	Indels 3;	Gaps 1;
Qy	47	ccaaagagaaattctgcagaaagtctgcatgaggcccaagcaagaaattactaaga	106		
Db	47	CCTGAGAAATTCATCCAGAGGTCAGGCCATGAAGATCTCTGACCACAAATGGGAGGA	106		
Qy	107	ggagttgccaatgaattcttgaagctgaaaaggcaatctaccaagtacaaagcgagacaa	166		
Db	107	CAACTTCGCCCGGGACTTCATCGGGTTAAGAAGATTGTCTACCAATATAGAACAGAAAA	166		
Qy	167	aacctactacaactgtggtgagaaatccaagaatatcaagaaaaacagataaaga	226		
Db	167	GATATATCCACAGCCACCTGGAGAAAAAGAAAAATGTTAAAAAGAACAGATACAAAGA	226		
Qy	227	tatttgccctatgattatagccgggtagaactatccctgataaacctctgtagaggttc	286		
Db	227	CATACCTGCCATTTGATCAGCGGAGTTAAATTGACATTAAGACATCTCTCAAGATTC	286		
Qy	287	cagctacataatcgcaactcattaaaggagtttatggaccacaaagccttatattgccac	346		
Db	287	AGATATATCAATGCAAAATTTTATAAAGGCGCTCTATGGCCAAAAAGCATATGTAGCAAC	346		
Qy	347	ccaggctcttatctacaacctctgcactctgagagatgattgggaatatagtt	406		
Db	347	TCAAGGACCTTTAGCAAAATACAGTATAGATTTTTCGGAGGATGATATGGAGTATAATGT	406		
Qy	407	ctttatcattgttgatgcatgagtgatgaaatgggaagaaaagtgtgagcgcta	466		
Db	407	TGFGATCATTTGTAATGGCTCGCGAGAAATTTGAGATGGGAGGAAAAAATGTGAGCGCTA	466		
Qy	467	ctgggtgagccaggagagatgcagctggaatttggtcccttctctgtatcctctgtaagc	526		
Db	467	TTGGCCTTTGTATGGAGAGAGACCCCAATAACGTTTGACCATTTAAAAATTTCTTGTGAGGA	526		
Qy	527	tgaaaaaggaatctgattataatacaggactctaaagttaagttaactagtgaaac	586		
Db	527	TGAACAAGCAAGAACAGACTACTTCAATCAGGACACTCTTACTTGAATTTCAAATGAATC	586		
Qy	587	tcgaactatctaccagtttctattacaagaattggccagaccatgatgtaccttcatctat	646		
Db	587	TCGTAGGCTGTATCAGTTTCATTTATGTGAAC TGCCAGACCATGATGATTCCTTTTCATCAT	646		
Qy	647	agacctattcttgactcaatctggatgtacgttgtttaccagaagagatgacagtgctcc	706		
Db	647	TGATTTCTTCTGGACATGATAGCTTAATAGAGAAATATATCAAGAACATCAAGATGTTCC	706		
Qy	707	catatcattcactgcagtgctgctggtggaagactggtgtattgtgctattgtga	766		
Db	707	TATTTGATTTCATTTGAGTGCAGGCGTGTGGAAGAACAGGTGCCATTTTGCCCA --TAGA	763		
Qy	767	ttatcatggatgtgtctaaaaagatgggataatctctgagaacttcagtggttttcagtt	826		
Db	764	TTATACGTGGAATTTACTAAAGCTGGAAAAATACCAGAGGAATTTTAATGTATTTAATTT	823		
Qy	827	gatccggggaatcggacacagagcgcttcatttagttccaacgcaggaacaatatgaact	886		
Db	824	AATACAGAAATGAGAACCAAAAGCATTTCTGCAGTACAAACAAAGGAGCAATATGAAC	883		
Qy	887	ggctcaaatgctgtattagaactatttaagagacagat	925		
Db	884	TGTTATAGAGCTATTGCCCCAACTGTTTGNAAAACAGCT	922		

RESULT 11
HUMPTPG1

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS

SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE

JOURNAL
MEDLINE

REFERENCE
AUTHORS

TITLE
JOURNAL

COMMENT

FEATURES

source

CDS

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 572;

Conservative 0;

Mismatches 304;

Indels 3;

Gaps 1;

Qy 47

ccaaagagaaattctgcagaaagtctgcatgaggcccaagcaagaaattactaaga

1004 a 556 c 577 g 801 t

PTFSKREQISENPTEADIGFNGRCRPGKPRDPSPSEWT"

Score 376.6;

DB 9;

Length 2938;

Pred. No. 2.8e-76;

Mismatches 304;

Indels 3;

Gaps 1;

Qy 47

Db	37	CCTGAGGAAATTCATCCACAGGGTCCAGGCGCATGAAGAGTCTCTGACCACAATGGGGAGGA	96
Qy	107	9gagtttgcccaatgaattcttgaagctgaagagcaatctaccaagtacaaggcagacaa	166
Db	97	CAACTTCGCCGGGACTTCATGCGGTTAAGAAGATTGCTTACCAAAATATAGAACAGAAAA	156
Qy	167	aacctatctacaactgtgctgagaatgccagaatatcaagaaaaacagataataagga	226
Db	157	GATATATCCACAGCCACTGGAGAAAAAGAAAAATGTTAAAAAGAACAGATACAAGGA	216
Qy	227	tattttgccttatgattagccgggtagaactatccctgataacctctgatgaggattc	286
Db	217	CATACTGCCATTTCATCACAGCCGAGTTAAATTGCACATTAAAGACTCTCTTCACAAGATTC	276
Qy	287	cagctacaataatgccaaacttcataagaggagtttatggacccaaggcttatattgcccac	346
Db	277	AGACTATATCAATGCAAAATTTTATPAAAGGCGCTGTATGGCCGCAAAAGCATATGTAGCAAC	336
Qy	347	ccagggtcccttatctacaacctccctggactctggagatgatttgggaatatagtgt	406
Db	337	TCAAGGACCTTTACCAAAATACAGTAATAGATTTTGTGGAGATGGTATGGGAGTATAATGT	396
Qy	407	ccctatcattgttatggcatgcatggagtagtgaatggaaagaaaaaagtgtgagcgcta	466
Db	397	TGTGATCATTTGTAATGCCTGCCGAGAAATTTGAGATGGGAAGGAAAAAATGTGAGCGCTA	456
Qy	467	ctgggctgagccagagagatgagctggaaatttggcccttctctgtatccgtgtgaagc	526
Db	457	TTGGCCCTTGTATGGGAGAGACCCCATTAAGCTTTGACCACTTTAAAAATTTCTTGTGAGGA	516
Qy	527	tgaaaaaaggaactctgattataatcaggacctctaaaagtttaagttcaatgtagaac	586
Db	517	TGAACAAGCAGACAGACTTTCATCAGGACACTCTTACTTCAATTTCAAAATGTAATC	576
Qy	587	tcgaactatctaccagtttcoattacaagaattggccagaccatgatgtaaccttcattat	646
Db	577	TCGTGAGCTGTATCAGTTTCATTATGTGTAAGTGGCCAGACCATGATGCTTCTTCATCATTT	636
Qy	647	agaccctattcttgagctcatctgggtagtgcgtgtgtaccagagatgacagagtctcc	706
Db	637	TGATTCATTCAGACATGATTAAGCTTAATCAGGAATATTCAGAACATCAAGATGTTCC	696
Qy	707	catatgcattcaatcagtcagtgctggagagactgtgttatttggctattgttga	766
Db	697	TATTTGTATTTCATTCGAGTCGAGGCTGTGGAAGAACAGGTGCCATTTGTGCCA---TAGA	753
Qy	767	ttatacatggatgttgctaaaaagatgggataaattccctgagaacttcagtttttcagttt	826
Db	754	TTATACGTGGAAATTTACTAAAAAGCTGGGAAATAACCCAGAGAGGAATTTAATGTATTAAATTT	813
Qy	827	gatccgggaaatgcgacacagagcgcttcattagtttcaaacgcaggacaacatgatgaact	886
Db	814	AATACAAGAAATGAGACACAAAGGCATTTGTCAGTACAAACAAGGAGGCATAATGAACT	873
Qy	887	ggctacaatgtcgttatagaactatttaagagacagat	925
Db	874	TGTTTCATAGAGCTATTGCCCAACTGTGTTGAAAAACAGCT	912

RESULT 12

G18337	G18337	2938 bp	DNA	linear	STS 28-SEP-1998
LOCUS	SWSS1985	Eric D. Green	Homo sapiens	STS genomic, sequence tagged site.	
ACCESSION	G18337				
VERSION	G18337.1	GI:1222794			
KEYWORDS	STS:				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 2938)				
AUTHORS	Bouffard, G.G., Iyer, L.M., Idol, J.R., Braden, V.V., Cunningham, A.F.,				

TITLE Weintraub, L.A., Mohr-Tidwell, R.M., Peluso, D.C., Fulton, R.S.,
 Leckie, M.P. and Green, E.D.
 JOURNAL A collection of 1814 human chromosome 7-specific STS
 MEDLINE Genome Res. 7 (1), 59-64 (1997)
 PUBMED 97189344
 REFERENCE 9037602
 AUTHORS 2 (bases 1 to 2938)
 TITLE Green, E.D.
 JOURNAL Human chromosome 7 STS (1997)
 COMMENT Unpublished
 Synonyms: PTPN12
 GDB: GDB:3754375
 GDB_DSEG: PTPN12
 Contact: Eric D. Green
 Genome Technology Branch
 National Human Genome Research Institute/NIH
 49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892
 Tel: 3014020201
 Fax: 3014024735
 Email: egreen@hgrl.nih.gov
 Primer A: GTAAGAATTTCTGCATGG
 Primer B: TTTCTAAAACTCCAGG
 STS size: 73
 PCR Profile:
 Presoak: 94 degrees C for 2.50 minute(s)
 Denaturation: 94 degrees C for 0.50 minute(s)
 Annealing: 45 degrees C for 0.75 minute(s)
 Polymerization: 72 degrees C for 0.75 minute(s)
 PCR Cycles: 35
 Thermal Cycler: PerkinElmer 9600
 Protocol:
 Template: 30-100 ng
 Primers: each 2 uM
 dNTPs: each 100 uM
 Taq Polymerase: 0.05 units/ul
 Total Vol: 10 ul
 Buffer:
 MgCl2: 1.5 mM
 KCl: 100 mM
 Tris-HCl: 10 mM
 NH4Cl: 5 mM
 pH: 8.6

Query Match	16.0%	Score 376.6;	DB 11;	Length 2938;
Best Local Similarity	65.1%;	Pred. NO. 2.8e-76;		
Matches 572;	Conservative	0;	Mismatches 304;	Indels 3; Gaps 1;
Qy	47	ccaaagagaaattctgcgaagttctctggatgagcccaagagaaaattactaaaga	106	
Db	37	ccttgaggaaattcattccagaggggtccaggccattgaagagttctgtaccacaaatggggagga	96	

QY	107	ggagtttgcaaatgaatttctgaagctgaaaggcaatctaccaagtacaaagcgacagaa	166
Db	97	CAACTTCGCGCGGACTTCATCGGTTAAGAAGATTGCTACCAATATAGAACAAGAA	156
QY	167	aacctatctacaactgctgaggaatccaaagaatacaagaagaaacagataaagga	226
Db	157	GATATATCCACAGCCACCTGGAGAAAAGAAGAAATTTAAAAAGACAGATACAAGGA	216
QY	227	tatttgccctatgattatagcgggtagaactatccctgataaacctctgatgagattc	286
Db	217	CATACCTGCCATTGATCAGCGGAGTTAAATTGACATTAAGACATTCCTTCACAGATT	276
QY	287	cagctacatcaatgccaactcattaaaggagttttgagcccaagcgttatattgccac	346
Db	277	AGACTATATCAATGCAAAATTTATAAAGGCGCTCTATGGCCAAAGACATATGTAGCAAC	336
QY	347	caaggctcctttatctacaacccctcctgacctctctgagagtgattgggaatatagtgc	406
Db	337	TCAAGGACCTTTAGCAAAATACAGTAATAGATTTTGGAGATGGTATGGGAGTATAATGT	396
QY	407	ccttatcatgttatggcagcatgagtatgaaatgggaagaaagaaagtgtgagcgcta	466
Db	397	TGTGATCATGTATNGCCTCGGAGAAATTTGAGATGGGAAGAAATATGTGAGCGCTA	456
QY	467	ctgggctgagccaggagagatgcagctggaatttggccctctctctgtatcctctgaaagc	526
Db	457	TTGCGCTTTGTATGGAGAGACCCCAATAAGTTTGACCATTTAAATTTCTTGTGAGGA	516
QY	527	tgaaaaaagaaatctgatatataatcsggacctcaaaagttaaagtcaatagtgaaac	586
Db	517	TGAACAAGCAAGAACACAGACTACTTTCATCAGGACACTCTTACTTTGAATTTCAAAATGAATC	576
QY	587	tgaactatctaccagtttattacaagaattggccagaccatgatgtaaccttcatctat	646
Db	577	TCTAGGCTGTATCAGTTTCATTTATGTGAAGTGGCCAGACCATGATGTTCCITTCATCAT	636
QY	647	agacctattctgagctcatctggagtgatgaactgtttaccagagagatgacagtgttcc	706
Db	637	TGATTTCTATCTGGACATGATAAGCTTAATGAGAGAAATATCAAGAACATGAAGATGTTCC	696
QY	707	catatgactcaactgagctgctgctgctggaaggagctgtgtatttctgctattgttga	766
Db	697	TATTTCTATTCATTTGAGTGTGAGGCTGTGGAAGAACAGGTGCCATTTGTGCCA---TAGA	753
QY	767	ttatcatgagtgctgctaaaagatggataattcctgagaacttcagtggtttcagttt	826
Db	754	TTATACGTGGAATTTACTAAAGCTGGGAAATATACCAGAGGAATTTAATGATTTAATTT	813
QY	827	gatccgggaaatcgacacagagcgcttcattagtttcaaacgcaggaacaatatgaact	886
Db	814	AATACAAGAAATGAGAACAACAAGGCATTCGCGAGTACAAACAAAGAGCAATATCAACT	873
QY	887	ggtctacaatgctgtattagaactatttaagagacagat	925
Db	874	TGTTATAGAGCTATTGGCCCAACTGTTTGAANAACAGCT	912
RESULT 13			
MMPTPPES			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			

JOURNAL	Biochem. J.	308 (Pt 2),	425-432	(1995)
MEDLINE	95289971			
REFERENCE	2 (bases 1 to 2676)			
AUTHORS	Tremblay, M.L.			
TITLE	Direct Submission			
JOURNAL	Submitted (01-MAY-1995)	M.L. Tremblay,	McGill University,	Dept of
COMMENT	Biochemistry,	3655 Drummond,	Montreal,	Quebec H3G 1Y6, CANADA
FEATURES	Related sequences:	X63440 and	S36169.	
Location/Qualifiers				
source	1..2676			
	/organism="Mus musculus"			
	/strain="BALB/c"			
	/db_xref="taxon:10090"			
	/chromosome="5"			
	/map="region A3-B"			
	/clone_lib="Lambda Zap II"			
	/dev_stage="17.5 days"			
1..14				
misc_feature	/note="synthetic linker"			
gene	330..2657			
	/gene="mptp-pest"			
CDS	330..2657			
	/gene="mptp-pest"			
	/codon_start=1			
	/product="protein-tyrosine-phosphatase"			
	/protein_id="CAA60477.1"			
	/db_xref="GI:804998"			
	/db_xref="MGD:MGI:104673"			
	/db_xref="SWISS-PROT:P35831"			
	/translation="MEQVEILRRFIQRVOAMKSPDHNGEDNFARFMRRLRLSTKYRT			
	EKIYPTATGKEENVKYKIDILFDHSRVKLTLPSTQSDSYDNANFKVGYGPKA			
	VATOGPFRTVIDFWIMWYVNVIMACREFEMGRKKERYCPLYGEDITTFAPF			
	KISCNEQARTDYERTLLLEFQNSRRLYQHYVNPDPDVPSSFDSTLDMISLMRK			
	YQEHEDVPICHCSAGCGRTGAICAIDYTNLLKAGKIPPEENVNLTQEMTQHSIA			
	VOTKQYELVHRAIAQLFEKQLQIEHGAQIRGNEITTCITVMSDDSDRYHKPVL			
	KPRTRSCLVEGDAKEELQPEPEHPVPIITPSPSAFTVTVTVMQSDSDRYHKPVL			
	HNASPEQPADLNRSYSDQWKSESAIEHIDKLKRNLSFEIKKVPLOEGPKSFD			
	GNTLLNRGHAIKIKSASSVDRTSKQELKSAGALKVDDVSONSCADCSAAHSHRAAE			
	SSEESQNSHTPPRDCPLDKKHVWTSLHGPENATVPVSDPKSNDHNSOTLKTV			
	SSTPSTAEAEADLDEHNSSPLIKAPLSFTNPLHSDDDHSDGSGDGAVTRNKTSI			
	STASATVSPASSAESACHRRVLPMSIAQEVAGCTPHSNAEKADADYSESPPLPERTP			
	ESFVLADMPVRPEWHLPENQSESEGLTSGNEKHADGIGIITEASADSPPAFSD			
	KDQDTTKSPAETVDFGNRQKPKGPRPPSEMT"			
BASE COUNT	774 a	632 c	682 g	588 t
ORIGIN				
Query Match	15.7%;	Score 370.2;	DB 10;	Length 2676;
Best Local Similarity	64.9%;	Pred. No. 8.3e-75;		
Matches	587;	Conservative	0;	Mismatches 303;
			Indels	15;
			Gaps	2;
QY	33	ctctgcagcatggacaaagaaattctgcagaattcctgcgattgagggcccaagcaag	92	
Db	321	CGCGGAGGATGGAGCAAGTGGAGATCCTGAGGAGGTTTCATCCAGAGGGTCCAGGCATG	380	
QY	93	aaattac-----taaaagagaggttgcgaatgaattctgaagctgaaaaag	140	
Db	381	AGAGTCCGATCACAATGGGAGGACAACTTCGCCCGGACTTCATGCGATTGAGAAGA	440	
QY	141	caatctaccaagtacaaaggcagacaaacctatctcactgaactggtgctgagaatgccaa	200	
Db	441	TTGTCTACCAAAATAGAACAGAAAAGATTATCCCAAGCCACTGAGAAAAAGAGAA	500	
QY	201	aatatcaagaaaaacagatataaggattatttgcctcatgattatagccgggtagaacta	260	
Db	501	AATGTTAAAGAACAGATATAAGGACATACTCCCATTTGATCACAGCCCGAGTTAAGTTG	560	
QY	261	tccttgataacctctgatgaggtatccagctacatcaatgccaaacttcattagggaggt	320	
Db	561	ACTTTGAAGACTCCATCCCAAGATTGAGATTATATCAATGCAATTTTATTAGGGGTG	620	
QY	321	tatggacccaaggcttatattgccaccagggttcctttatctacaaacctctcgtgacttc	380	

Db 846 ACAGTGCTATTGTGGCCA--TCGACTACACGTGGAACTTACTGAAAGCAGGGAATAAT 902
Qy 801 cctgagaactcagtggtttcagttgacccggaataatcgacacagagcccttcatta 860
Db 903 CCAGAGGAATTAATGTATTAATTAACAAGAAATGAGAACAACAAGGCACCTCTGCA 962
Qy 861 gttcaaacgcaggaacaatatgaactggctctacaaatgctgtattagaactatttaagaga 920
Db 963 GTACAAACGAAGGACGATGATGAACCTCGTCCATAGAGCTATTGCCCCAAGCTGTTGAANA 1022
Qy 921 cagat 925
Db 1023 CAGCT 1027
RESULT 15
AX306312 AX306312 2983 bp DNA linear PAT 11-DEC-2001
LOCUS Sequence 1063 from Patent WO0188188.
ACCESSION AX306312
VERSION AX306312.1 GI:17645552
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (sites)
AUTHORS Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 1063 22-NOV-2001;
School Juridical Person Nihon University (JP)
FEATURES
Location/Qualifiers
source 1..2983
/organism="Mus musculus"
/db_xref="taxon:10090"
BASE COUNT 923 a 634 c 681 g 745 t
ORIGIN
Query Match 15.6%; Score 368; DB 6; Length 2983;
Best Local Similarity 64.9%; Pred. No. 2.6e-74;
Matches 583; Conservative 0; Mismatches 300; Indels 15; Gaps 2;
Qy 33 cctctcagcatggaccacaagagaaaattctgcagaagtctcctggatgagggcccaagcaag 92
Db 43 CGCGGGAGGATGGAGCAAGTGGAGATCTCTGAGGAGGTTCTCCAGAGGGTCCAGGCCATG 102
Qy 93 aaaaatac-----taagagagagtttgccaatgaaattctcgaagctgaaaagg 140
Db 103 AAGAGTCCGGATCACAATGGGGAGGACAACTTCGCCCGGGAGCTTCATGCGATTGAGAA 162
Qy 141 caatctacaagtcagaagcagacaaaacctctatctacaactgtgctgagaatcccaag 200
Db 163 TTGCTTACCAATATAGAACAGAAAGATTTATCCACAGCCACTGGAGAAAAAGAA 222
Qy 201 aatatcaagaaaaacagatatgaagatatatttgccctatgattatagccgggtagaacta 260
Db 223 AATGTTAAAAGACAGATATAAGGACATACCTGCCATTGTATCAGCCGAGTTAAGTTG 282
Qy 261 tccctgataacctctgatgaggtatccagctacatacaatgccaaacttcattaaaggagtt 320
Db 283 ACTTTGAAGACTCCATCCCAAGATTTCAGATTATATCAATGCAAAATTTATTAAAGGTGTG 342
Qy 321 tatggaccaagcatttatattgccaccaagggtcctttatctacaacccctcctggacttc 380
Db 343 TATGGGCCAAAAGCATATGTGGCAACCCCAAGGGCCCTTTCCGGAATACAGTCTACAGACTTC 402
Qy 381 tggagatgatttgggaatacagttcctctatctattgttattgctgcatgaggtatgaa 440
Db 403 TGGAGGATGATATGGAGTATTAATGTTGTGATGATCGTGATGGCCTGTCGAGAAATTTGAG 462
Qy 441 atgggaaagaaaaagtgtgagcgctactgggctgagccaggagagatgcagctggaattt 500

Db 463 ATGGGAAGAAAAAGTGTGAGCGCTACTTGGCTTTTGTATGGAGAAGATCCTTATAACATTT 522
Qy 501 ggccttctctgtatcctctgtaagctgaaaaaaggaaactctgattatataatcaggagact 560
Db 523 GCACCATTTAAAATTTCTTGTGAAANTGAACAAGCAAGCAACCGACTACTTCATCCGAACA 582
Qy 561 ctaaaagttaagttcaatagtagaaactcgaactatctaccagttttcattacaagaattgg 620
Db 583 CTTTACTTGAATTCAAAATGAATCCCGTCGGCTCTATCATCAGTTTCATTACGTGAACCTGG 642
Qy 621 ccagaccatgatgtacctctcatatagaccctattcttgagctcatctctggatgagctg 680
Db 643 CCAGACCATGATGTTCCCTTCGTCATTTGATTCTATTCTGGACATGATTAAGCTTAATGAGG 702
Qy 681 tgttaccagaaggatgacagtggttcccatatgcttcacactcagtgctgctgctggaagg 740
Db 703 AAATACCAAGAACATGAAGATGTGCCTATTGTATTCATTGACGTGCTGGCTGTGGACGA 762
Qy 741 actggtgttattgtgtattgttattgattatcacatggtattgctaaaaagatgggataatt 800
Db 763 ACAGGTGCTATTGTGCCA---TAGATTACACGTGGAACCTTACTGAAAGCAGGGAATAAT 819
Qy 801 cctgagaactcagtggttttccagtttgatccgggaataatcgggacacagagcccttcatta 860
Db 820 CCAGAGGAATTTAATGTATTTAATTTAATACAAGAAATGAGAACAACAGAGGCACCTCGCA 879
Qy 861 gtccaacgcgaggaacaataatgaactggtctacaaatgctgtattagaactatttaaga 918
Db 880 GTACAAACAAGGAGCAGTATGAACCTTGTTCATAGGGCTATTGCTCAACTGTTTGAAA 937

Search completed: May 26, 2002, 10:35:23
Job time: 12938 sec

QY 632 AGEFSNPVKSLSSAVKVIKIGTSLWEG-----GTSEPKKFDSDVILRPSKSVKLRSPKSG 686
Db 691 -----HNTP-----VRSSELSQSERSEKQSEG--LITSEKCDHPAGG 730
QY 687 KNF 689
Db 731 IHY 733
RESULT 3
S55345
protein-tyrosine-phosphatase (EC 3.1.3.48) PTP-PEST - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Oct-1995 #sequence_revision 01-Nov-1996 #text_change 22-Jun-1999
C:Accession: S55345; S54261; I48666
R:Charest, A.; Wagner, J.; Shen, S.H.; Tremblay, M.L.
Biochem. J. 308, 423-432, 1995
A:Title: Murine protein tyrosine phosphatase-PEST, a stable cytosolic protein tyrosine P
A:Reference number: I48666; MUID:95289971
A:Accession: S55345
A:Molecule type: DNA
A:Residues: 1-775 <CHA>
A:Cross-references: EMBL:X86781; NID:g804997; PIDN:CAA60477.1; PID:g804998
R:Charest, A.; Wagner, J.; Shen, S.H.; Tremblay, M.L.
submitted to the EMBL Data Library, May 1995
A:Description: Murine protein tyrosine phosphatase-PEST, a stable cytosolic protein tyro
A:Reference number: S54261
A:Accession: S54261
A:Molecule type: DNA
A:Residues: 1-127,'M',129-309,'R',311-775 <CHW>
A:Cross-references: EMBL:X86781; NID:g804997; PIDN:CAA60477.1; PID:g804998
A:Experimental source: strain BALB/c
C:Genetics:
A:Map position: 5A3-B
C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 12; protein-tyrosine-phosph
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatas
F:58-282/Domain: protein-tyrosine-phosphatase homology <PTP>
F:231/Active site: Cys (phosphocysteine intermediate) #status predicted
F:237/Binding site: substrate phosphate (Arg) #status predicted
Query Match 31.2%; Score 1127.5; DB 2; Length 775;
Best Local Similarity 36.6%; Pred. No. 1.9e-62;
Matches 280; Conservative 109; Mismatches 166; Indels 209; Gaps 28;
QY 1 MQOREILOKFLDEAQSCKIT----KSEFANEFLKLRQSTKYKADKYPTTVAENAKNIK 56
Db 1 MEQVEILRRFIQRVQAMKSPDHNGEDNFARDFMLRRLSTKTKYTKIYPTATGEKEENVK 60
QY 57 KNYKDIPLDYSRVELSLITSDSSYINANFIKGVGPKAYATQGPLSTLLDFWRM 116
Db 61 KNYKDIPLFDHSRVKLTLPQSDSDYINANFIKGVGPKAYATQGPFRNTVIDFWRM 120
QY 117 IWYSYLIIVMACMEYMGKKKERYWAEFPGEMQLEFGFPFSVCEAEKRKSDYIIRTLKV 176
Db 121 IWEYNVVIIVMACREFEMGRKCKERYWPLYGEDPITFAPEFKISCENEQARTDYFIRTLLL 180
QY 177 KFNSERTIYQFYHKNPDHVPSSDPILELIWDVRCQEDDSVPCIHCSAGCGRTGV 236
Db 181 EFQNESRRLYQFYHKNPDHVPSSDPILELIWDVRCQEDDSVPCIHCSAGCGRTGA 240
QY 237 ICAIVDYTWMLLKDGIIPEVNFVSLIREMRTORPSLVOTEOYELVYNVAVLELFXKROMD 296
Db 241 ICAI-DYTNLLKAGKIPEEFNVFNLIQEMRTORHSAVQTKQYELVHRIAQLFKEQLQJ 299
QY 297 V-----IRDKHSGTESQAKCIPEKNHNTLOADYSNLPKSTTKAAKMMNQRTKME 348
Db 300 LYBTHGAQKADNCEITTTGTVMWSSIDSE-----KQDSPPPKPPPTRS-----CLVEGDAKEE 351
QY 349 IKE-----SSSDFPFRS-----EISAKEELVLPKAKSTSTF 379
Db 352 ILQPPPEHPVPPIITLTPSPSFAFTVTVTVMQDSDRYHPKPYLHMASPEQ----HPA----- 402

QY 380 DFLELNYSFQKADNADTTMKW-----OTKAPPIVGEPLQK-HQSLDLGLSLLF 423
Db 403 ----DLNRSYDKSAD---QWCKSESAIEHIDKKLERNLSEIKKVPLOEGPKSGFDGNTLLN 456
QY 424 EGCSNKPVNAAGRYFNYSKVPITRTKSTPPEL-----IQORETKEV----- 454
Db 457 RG--HAIKIKSA-----SSSVVDRT--SKQELSGAGALKVDDVDSQNSCADCSAAHSHRAAE 508
QY 465 DSKENESYLESDPHSCFVEMQAQKVMHYSSAELNYSL-----PYDSK 507
Db 509 SSEESQSNHTPPRPDC---LPLDKKGHVT-----WSLHGPENATVPVPSDCKGSPDNHS 560
QY 508 HQLRNASN-----VKHDDSSALGVSYIPLVENP-----YFSSWPPSGTS 547
Db 561 QTLKTVSSTPNSTAEAEADLTEHNSS-----PLLKAPLSFTNPLHSDDDWHSDDGGS 612
QY 548 SKMSLDLPEKQDQTVFPSSLLPTSTSLFSYNSHSLNSLSPNTSSLLNQSSAV----- 603
Db 613 S-----DCAV-----TRNKTSISTASAT-VSPASSAESACHRRV 645
QY 604 -----LATAPR-----IDDEIPPLPVRTPESEFIVVE 630
Db 646 LPMISARQEVAGTPHSGAEKADADVSESPPLPERTPESEFVLAD 689
RESULT 4
JH0609
protein-tyrosine-phosphatase (EC 3.1.3.48) P19 - mouse
N:Alternate names: protein-tyrosine-phosphatase PPTTY43
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: JH0609; PS0365; PS0369; PS0366; G61180
R:den Hertog, J.; Pals, C.E.G.M.; Jonk, L.J.C.; Kruijer, W.
Biochem. Biophys. Res. Commun. 184, 1241-1249, 1992
A:Title: Differential expression of a novel murine non-receptor protein tyrosine phos
A:Reference number: JH0609; MUID:92272714
A:Accession: JH0609
A:Molecule type: mRNA
A:Residues: 1-773 <DEN>
A:Cross-references: GB:X63440; GB:S36169; NID:g416181; PIDN:CAA45037.1; PID:g416182
A:Experimental source: embryonic carcinoma cell, P19 cell
A:Accession: PS0365
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 88-91,'G',93-110,'G',112-118,'S',120,'T',122 <DE2>
A:Experimental source: embryonic carcinoma cell, P19 cell, clone PTP33
A:Accession: PS0369
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 88-91,'G',93-109,'LG',112-120,'T',122 <DE3>
A:Experimental source: embryonic carcinoma cell, P19 cell, clone PTP59
A:Accession: PS0366
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 88-91,'KY',94-109,'LA',112-118,'S',120-122 <DE4>
A:Experimental source: embryonic carcinoma cell, P19 cell, clone PTP42
R:Yi, T.; Cleveland, J.L.; Ihle, J.N.
Blood 78, 2222-2228, 1991
A:Title: Identification of novel protein tyrosine phosphatases of hematopoietic cells
A:Reference number: A61180; MUID:92032882
A:Accession: G61180
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 124-127,'I',129-229 <YIA>
C:Comment: This protein is located in the cytoplasm.
C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 12; protein-tyrosine-ph
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosph
F:55-299/Domain: phosphatase catalytic domain #status predicted <PCD>
F:58-282/Domain: protein-tyrosine-phosphatase homology <PTP>
F:231/Active site: Cys (phosphocysteine intermediate) #status predicted
F:237/Binding site: substrate phosphate (Arg) #status predicted

Query Match 30.6%; Score 1104.5; DB 1; Length 773;
Best Local Similarity 35.6%; Pred. No. 5e-61;
Matches 273; Conservative 103; Mismatches 175; Indels 215; Gaps 26;

QY 1 MDQREILQKFLDEAQSCKIT----KEEFANFELKLRKSTYKADKTYPTTVAENAKNIK 56
DB 1 MEQVEILRRFQORVOAMKSPDHNGEDNFARDFMRRLRSTKYRTKEIYPTATGEKEENVK 60
QY 57 KNRKIDILPYDSRVVELSLITSDSSYINANFIKGVYGPAYIATQGPLSTLLDFWRM 116
DB 61 KNRKIDILPFDHSRVKLLTKTPSQSDSYINANFIKGVYGPAYVATQGFRTNVIDFWRM 120
QY 117 IWEYSVLIIVMACMEYEMKKKERYWAPGEMQLEFQFVSVCBAEKRSYIIRTLLK 176
DB 121 IWEYNVMIIVMACREFEMGRKKERYWPLYGEDPITFAFKTSCNEQARTDYFIRTL 180
QY 177 KFNSETRIYQFHYKNWPDHVPSSIDPILFELIWDVRCYQEDDSVPICHSAGCGRTGV 236
DB 181 EFQESRRLYQFHYVNPDPHVPSSFDLSILDMISLMRYQEHEDVPICHSAGCGRTGA 240
QY 237 ICAIVDYTWMLLKDGIIIPENFVSFLIREMRTQPSLVQTOEQYELVYNVLELFRK 296
DB 241 ICAI-DYTWMLLKAGKIPFEEFNVLQEMRTQHSVQTKEQYELVHRAIAQLPEK 299
QY 297 VI-----RDKHS-----GTSQAKHCPIE 315
DB 300 CMKMEHRRSVVMKPLLELWVLIARRDLAKAAADSKLPCRRGCGQGRNTATRTSPG 359
QY 316 KNTLQADSYSNLPKSTTKAKM-----MNQORTKWEIKESSPDRFISE 361
DB 360 AHPDAITSFS--LPNVTVTCRTVTGTQSQCCWPHOSNTQPTSTE-----AMINORT-- 411
QY 362 ISAKEELVLHPAKSTSFDFELNYSFKDNADTTMKWOTKAPPIVGEPLQKHQSLDGL 421
DB 412 -----NGAKSESALEHI-----DKLERNLSFELKKVPLQEGP-----KSF 452
QY 422 LFEGCSNKPVNNAAGRYNSKVPITRTKSTPEL-----IQORETKEY----- 464
DB 453 LNRG--HAIKIKSA-----SSSVVDRT--SKQELSGALGVDDVDSQNSCADCSAAHSHRA 504
QY 465 --DSKENFYSLEQPHDSCFVEMQAQKVMHVYSSAELNYSL-----PYD 505
DB 505 ARSSESOSNSHTPPRPGC--LPDLKKGHVY----WSLHGHPENATVPVPSDQKSPDN 556
QY 506 SRHQIRNASN-----VKHHDSSALGVSYIPLVBNP-----YFSSWPPSPG 545
DB 557 HSQTLKTVSTPNSAEEAHDLTEHNS-----PLKAPLSFTNPLHSDDDHSDG 608
QY 546 TSKMSLDLPEKODGTVPFSSLLPTSSLSFYSYNSHSLNSPTNSSLNQSSAV-- 603
DB 609 GSS-----DGAV-----TRNKTSTISATAT--VSPASSAESACHR 641
QY 604 -----LATAPR-----IDDEIPLPPLVPTPESFIVVE 630
DB 642 RVLPMISIAQEVAGTHPSGAEKDADVSESPPLPPTPESFVLAD 687

RESULT 5
S48748
Protein-tyrosine-phosphatase (EC 3.1.3.48), probable nonreceptor type 12 splice form -
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: S48748
R:Moriyama, T.; Kawanishi, S.; Inoue, T.; Imai, E.; Kaneko, T.; Xia, C.; Takenaka, M.; N
FEBS Lett. 353: 305-308, 1994
A:Title: cDNA cloning of a cytosolic protein tyrosine phosphatase (RKPP) from rat kidne
A:Reference number: S48748; MUID:95046282
A:Accession: S48748
A:Molecule type: mRNA
A:Residues: 1-382 <NR0>
A:Cross-references: GB:D38072; NID:9567262; PIDN:BAA07266.1; PID:G699627
C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 12; protein-tyrosine-phosph
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatase

F:58-282/Domain: protein-tyrosine-phosphatase homology <PRP>
F:231/Active site: Cys (phosphocysteine intermediate) #status predicted
F:237/Binding site: substrate phosphate (Arg) #status predicted

Query Match 28.8%; Score 1042; DB 1; Length 382;
Best Local Similarity 53.5%; Pred. No. 1.4e-57;
Matches 201; Conservative 62; Mismatches 89; Indels 24; Gaps 5;
QY 1 MDQREILQKFLDEAQSCKIT----KEEFANFELKLRKSTYKADKTYPTTVAENAKNIK 56
DB 1 MEQVEILRRFQORVOAMKSPDHNGEDNFARDFMRRLRSTKYRTKEIYPTATGEKEENVK 60
QY 57 KNRKIDILPYDSRVVELSLITSDSSYINANFIKGVYGPAYIATQGPLSTLLDFWRM 116
DB 61 KNRKIDILPFDHSRVKLLTKTPSQSDSYINANFIKGVYGPAYVATQGPLANTVIDFWRM 120
QY 117 IWEYSVLIIVMACMEYEMKKKERYWAPGEMQLEFQFVSVCBAEKRSYIIRTLLK 176
DB 121 IWEYNVMIIVMACREFEMGRKKERYWPLYGEDPITFAFKTSCNEQARTDYFIRTL 180
QY 177 KFNSETRIYQFHYKNWPDHVPSSIDPILFELIWDVRCYQEDDSVPICHSAGCGRTGV 236
DB 181 EFQESRRLYQFHYVNPDPHVPSSFDLSILDMISLMRYQEHEDVPICHSAGCGRTGA 240
QY 237 ICAIVDYTWMLLKDGIIIPENFVSFLIREMRTQPSLVQTOEQYELVYNVLELFRK 296
DB 241 ICAI-DYTWMLLKAGKIPFEEFNVLQEMRTQHSVQTKEQYELVHRAIAQLPEK 299
QY 297 VIRKHSSTESQAKHCIPKNTLQADSYSNLPKSTTKAAKMMNQORTKWEIKESS 356
DB 300 -LYETHGAKITDGNFSTGNMVSIDSEKQSDPPP-----KPPRTRSLCLVEG---- 346
QY 357 FRTSISAKEELVLP 372
DB 347 -----DAKEELQPP 356

RESULT 6
A46546
leukocyte common antigen long splice form precursor - human
N:Alternate names: C045; protein-tyrosine-phosphatase, receptor type c; T200 glycopro
N:Contents: leukocyte common antigen intermediate splice form; leukocyte common anti
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jan-2000
C:Accession: A46546; B46546; A29449; B29449; I57658
R:Streuli, M.; Hall, L.R.; Saga, Y.; Schlossman, S.F.; Saito, H.
J. Exp. Med. 166, 1548-1566, 1987
A:Title: Differential usage of three exons generates at least five different mRNAs en
A:Reference number: A46546; MUID:88061067
A:Accession: A46546
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1304 <STR>
A:Cross-references: GB:Y00638
A:Experimental source: clone LCA.6/2
A:Accession: B46546
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-32,99-264 <ST2>
A:Cross-references: GB:Y00638
A:Experimental source: clone LCA.111 and clone LCA.260
A:Accession: C46546
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-31,193-264 <ST3>
A:Cross-references: GB:Y00638
A:Experimental source: clone LCA.1
R:Ralph, S.J.; Thomas, M.L.; Morton, C.C.; Trowbridge, I.S.
EMBO J. 6, 1251-1257, 1987
A:Title: Structural variants of human T200 glycoprotein (leukocyte-common antigen).
A:Reference number: A91066; MUID:87275816
A:Accession: A29449

A:Molecule type: mRNA
A:Residues: 1-31,193-649,'L',651-869,'G',871-872,'A',874-1206,'P',1208-1304 <RAL>
A:Cross-references: GB:Y00062; NID:q34275; PIDN:CAA68269.1; PID:q34276
A:Experimental source: clones PHLC-1 and lambdaHLG1
A:Accession: B29449
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 32-192 <RA2>
A:Experimental source: clone HLC-2
R:Tsai, A.Y.; Streuli, M.; Saito, H.
Mol. Cell. Biol. 9, 4530-4555, 1989
A:Title: Integrity of the exon 6 sequence is essential for tissue-specific alternative splicing
A:Reference number: I57658; MUID:90066468
A:Accession: I57658
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 146-192 <RES>
A:Cross-references: GB:M29253; NID:gl87020; PIDN:AAA59497.1; PID:q553521
C:Genetics:
A:Gene: GDB:PTPRC; CD45
A:Cross-references: GDB:119768; OMIM:151460
A:Map position: lq31-lq32
C:Superfamily: leukocyte common antigen; leukocyte common antigen cytosolic domain homolog
C:Keywords: alternative splicing; glycoprotein; phosphoprotein; phosphoric monoester hyd
F:594-1235/Domain: leukocyte common antigen cytosolic domain homology <LAC>
F:575-899/Domain: protein-tyrosine-phosphatase homology <PTP>
F:851/Active site: Cys (phosphocysteine intermediate) #status predicted
F:857/Binding site: substrate phosphate (Arg) #status predicted

Query Match 14.2%; Score 514.5; DB 1; Length 1304;
Best Local Similarity 28.2%; Pred. No. 6.3e-24;
Matches 166; Conservative 86; Mismatches 241; Indels 95; Gaps 19;
QY 2 DQREILQK-----FLDQAQSKITKE--EPANFELKLRKSTIKYKADTKY 44
DB 612 EQGELVERDEKQLMNVPEITHADILLETYYKRIKIADEGRPLAEFQSPRVFSK-----F 665
QY 45 PTTVAENAKNIKKNRYKDILPYDSRVLSLTISDESSYINANFIKGVYGPAYATATQG 104
DB 666 PIKARKPFNONKNRVVDILPYDNRYVSEIDNGDAGSNYINASYIDGFEKPKYITAAQG 725
QY 105 PLSTLLDFWRMTWEXSVLIIVACMEYEMGKKKERYWAEPEGMOLFEGFVSVCBAEK 164
DB 726 PROETVDDFWRMTEQKATVIVWTRCEEGRNKKCAEYWFPMSEEGTAFQDVVVKINQHK 795
QY 165 RKSDYIIRTLKV---KFNSETRTYOFHYKNWPDHVPSSIDPILFELIWDVRCYQEDDSV 221
DB 786 RCPDYIIQKLVNKKKEKATGREVTHIQFTSWPDHGVDPDPLHLLKLRNRVNAFNSPFG 845
QY 222 PICIHCSAGCGRTGVCAYVDYTWMLLKDGIIPENFSVFLIREMRTORPSLVQTQOYE 281
DB 846 PIYVHCSAGVGRGTGIIGID-----AMLEDETENKVDVYGVYVVKLRQRCLMVQVEAQYI 901
QY 282 LVYNVAVLEL--FKRQMDVIRKHSGETSOAKHICPEKNHTLOADSYSPNLPK-STTKAAK 338
DB 902 LIHQALVEYVQFETEVNLSLHPLHNMKKRDPSPSPLEAE--FQRLPSTRSWRQTQ 959
QY 339 MMNQQRTKMEIKESSDFRTSEISAKEELV-----HPAKSSTFDLELNYSFQKNAD 393
DB 960 IGNOEENKSNRNSNVPIDYDNRVPLKHELEMSKESHDSDSDSDSDSEEPSESKYNAS 1019
QY 394 TTKW-WOTKAPPIVGEPLQKHSLDGLSLF-----BGCNSKSPVNA 434
DB 1020 FIMSYWPKPEYVIAAQGPL-KETIGDEFWQMIFQRKVIVMLTELKHLQDGEICAQ---YWG 1075
QY 435 AGR--YFNSKVPITRT--KSTPEEL--IQQRETKEVDSKENFSY-----LESQPHD- 479
DB 1076 EGKQTYGDIEVDLKDQDSTYTLRVFELRHSKRKDSRTYQYQYTNWSVEQLPAEPKEL 1135
QY 480 SCFVEMQAQKVMHYSSAELNYSFLYDSKHKQIRNASNVKHHDSALGVY 527
DB 1136 ISMIQVVKQKLPQKNSSEGN-----KHKSTPLLIH 1166

RESULT 7

S68700
HPTP beta-like tyrosine phosphatase precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C:Accession: S68700
R:Kuramochi, S.; Matsuda, S.; Matsuda, Y.; Saitoh, T.; Ohsugi, M.; Yamamoto, T.
FEBS Lett. 378, 7-14, 1996
A:Title: Molecular cloning and characterization of Byp, a murine receptor-type tyrosi
A:Reference number: S68700; MUID:96140699
A:Accession: S68700
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1238 <KUR>
A:Cross-references: GB:D45212; NID:gl208432; PIDN:BAA08146.1; PID:gl208433
C:Genetics:
A:Map position: 2E1-2
A:Superfamily: protein-tyrosine-phosphatase, receptor type J; fibronectin type III re
C:Keywords: phosphoprotein
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-1238/Product: HPTP beta-like tyrosine phosphatase #status predicted <MAT>
F:267-347/Domain: fibronectin type III repeat homology <3FR>
F:266-1188/Domain: protein-tyrosine-phosphatase homology <PTP>
F:1140/Active site: Cys (phosphocysteine intermediate) #status predicted
F:1146/Binding site: substrate phosphate (Arg) #status predicted

Query Match 14.2%; Score 513; DB 2; Length 1238;

Best Local Similarity 39.0%; Pred. No. 7.2e-24;
Matches 115; Conservative 50; Mismatches 110; Indels 20; Gaps 7;

QY 15 QSKKIKKEEFANFELKLRKST-----KYKADK-----TYPTVAENAKNIKKNRYKDILP 65
DB 918 KSKLIRVENFEAYFKKQADSNCGFAEEYEDLKLIGLSLKPYKTAETAEENKGNKNNVLP 977
QY 66 YDSRVLSLTISDESSYINANFIKGVYGPAYATATQGPLSTLLDFWRMWEYSVLII 125
DB 978 YDISRVLSVQTHSTD--DYINANMPGYHSHKDFIATQGPLPNTLKRDFWRMWEKNVYAI 1036
QY 126 VMACMEYEMGKKKERYWAEPEGMOLFEGFVSVCBAEKRSKYIIRTLKV--FNSETR 183
DB 1037 VMLTKCQVQGRTKCEEYW--PSKQADYGDITVAMTSEVVLPEWTIRDFVVKMNSHES 1094
QY 184 TIYOFHYKNWPDHVPSSIDPILFELIWDVRCY--QEDDSVPICIHCSAGCGRTGVCAYI 241
DB 1095 PLRQFHFTSWPDHGVDPDTHLLINFRYLVRDYMKQIPPEPILVHCSAGVGRGTGTFIAD 1154
QY 242 DYTWMLLKDGIIIPENFSVFLIREMRTORPSLVQTQOYELVYNVAVLELFRQMD 296
DB 1155 RLIIQIENENTV-----DVYGIYDLRMHRPLMVQTDQYVFLNQCVLDIIRAQND 1205

RESULT 8

I38670
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type J precursor - human
N:Alternate names: density enhanced phosphatase-1 (DEP-1); protein-tyrosine phosphata
C:Species: Homo sapiens (man)
C:Date: 01-Mar-1996 #sequence_revision 08-Mar-1996 #text_change 22-Jun-1999
C:Accession: I38670; I52599
R:Ostman, A.; Yang, Q.; Tonks, N.K.
Proc. Natl. Acad. Sci. U.S.A. 91, 9680-9684, 1994
A:Title: Expression of DEP-1, a receptor-like protein-tyrosine-phosphatase, is enhance
A:Reference number: I38670; MUID:95024024
A:Accession: I38670
A:Molecule type: mRNA
A:Residues: 1-1337 <RES>
A:Cross-references: EMBL:U10886; NID:g558754; PID:g558755
A:Experimental source: HeLa cells
R:Honda, H.; Inazawa, J.; Nishida, J.; Yazaki, Y.; Hirai, H.
Blood 84, 4186-4194, 1994
A:Title: Molecular cloning, characterization, and chromosomal localization of a novel

A:Reference number: I52599; MUID:95086212
A:Accession: I52599
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-216 'LTCVRKAA', 225-260 'G', 262-285 'GTEGLDASNTSERSRA', 302, 'S', 304, 'TAPVHDE
A:Cross-references: GB:D37781; NID:9633072; PIDN:BA07035.1; PID:9633073
C:Comment: Enhanced expression of this protein with increasing cell density suggests a r
C:Genetics:
A:Gene: GDB:PTPRJ
A:Cross-references: GDB:385040; OMIM:600925
A:Map position: 19q13.4-19q13.4
C:Function:
A:Description: catalyzes hydrolysis of peptidyl-phosphotyrosine to peptidyl-tyrosine and
C:Superfamily: protein-tyrosine-phosphatase, receptor type J; fibronectin type III repea
C:Keywords: glycoprotein; phosphoprotein; phosphoric monoester hydrolase; receptor; tran
F:1-35/Domain: signal sequence #status predicted <SIG>
F:36-133/Product: protein-tyrosine-phosphatase, receptor type J #status predicted <MAT>
F:118-197/Domain: fibronectin type III repeat homology <3FNA>
F:206-283/Domain: fibronectin type III repeat homology <3FNB>
F:284-356/Domain: fibronectin type III repeat homology #status atypical <3FNC>
F:365-445/Domain: fibronectin type III repeat homology <3FND>
F:453-530/Domain: fibronectin type III repeat homology <3FNE>
F:539-617/Domain: fibronectin type III repeat homology <3FNF>
F:720-804/Domain: fibronectin type III repeat homology <3FNG>
F:972-988/Domain: transmembrane #status predicted <TMN>
F:1065-1287/Domain: protein-tyrosine-phosphatase homology <PTP>
F:1272-82,93,104,142,172,192,231,258,278,342,351,376,391,396,413,431,501,525,536,582,603,6
F:1239/Active site: Cys (phosphocysteine intermediate) #status predicted
F:1245/Binding site: substrate phosphate (Arg) #status predicted

Query Match 14.1%; Score 508; DB 1; Length 1337;
Best Local Similarity 38.3%; Pred. No. 1.7e-23;
Matches 113; Conservative 51; Mismatches 111; Indels 20; Gaps 7;

QY 15 QSKKTKKEFAEFLLKQST-----KYKADK-----TYPTVAENAKNKNRYKDILP 65
DB 1017 KSKLIRVENFEAFYKQKQADSCNGFAEEYEDLKGISQPKYAAELAENRGKNRYNNVLP 1076

QY 66 YDYSRVELSLTSDSESSVINANFKGVGPKAYIATQGLSTLLDFWRMIWEYSVLII 125
DB 1077 YDISRVKLSVQTHSTD-DYINANYMPGYHKKDFATQGLPENTLKDFFRWVWEKRYVAI 1135

QY 126 VMACMEYEMGKKKERYWAEPMOLEFGPFVSCEAEKRSYDIIRTLKVK--FNSETR 183
DB 1136 IMLTKCVOGRTKCEYW--PSKAQDYGDITVAMTSEIVLPETIRDTFKNIQTSESH 1193

QY 184 TIYOHYKNWPHDVPSSIDPILLELWDVRCY--QEDDSVPICIHCSACCGRTGVCATV 241
DB 1194 PLRQHFHTSWPDHGVPTDITDLINFRYLVRDYMKQSPPEPILVHCSAGVGRGTGFIAID 1253

QY 242 DYTWMLLKGGIIPENFVSFLIREMRTQPSLVLTQEQYVLYNNAVLFLFKQMD 296
DB 1254 RLIIQIENENTV-----DVYGVIVDLMRHPLMVQVEDQYVFLNQCVLDIVRSQKD 1304

RESULT 9
A28334
protein-tyrosine-phosphatase (EC 3.1.3.48) Ly-5 precursor (B-cell variant) - mouse
N:Alternate names: 200K leukocyte common antigen; CD45; p19tyl; T-cell surface glycoprot
N:Contains: protein-tyrosine-phosphatase (T-cell variant)
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 21-Jan-2000
A:Accession: A28334; A29381; A61180; A60933; A33522; A29075; I54450
R:Thomas, M.L.; Reynolds, P.J.; Chain, A.; Ben-Neriah, Y.; Trowbridge, I.S.
Proc. Natl. Acad. Sci. U.S.A. 84, 5360-5363, 1987
A:Title: B-cell variant of mouse T200 (Ly-5): evidence for alternative mRNA splicing.
A:Reference number: A28334; MUID:87260986
A:Accession: A28334
A:Molecule type: mRNA
A:Residues: 1-1291 <THO>
A:Cross-references: GB:M22455
R:Saga, Y.; Tung, J.S.; Shen, F.W.; Boyse, E.A.

Proc. Natl. Acad. Sci. U.S.A. 83, 6940-6944, 1986
A:Title: Sequences of Ly-5 cDNA: Isoform-related diversity of Ly-5 mRNA.
A:Reference number: A29381; MUID:86313686
A:Accession: A29381
A:Molecule type: mRNA
A:Residues: 1-30,170-517, 'NNT', 521-527, 'G', 529-555, 'S', 557-587, 'S', 589-905, 'Q', 907-93
A:Cross-references: GB:M14342; NID:9198914; PIDN:AAA39458.1; PID:9198915
R:Yi, T.; Cleveland, J.L.; Ihle, J.N.
Blood 78, 2222-2228, 1991
A:Title: Identification of novel protein tyrosine phosphatases of hematopoietic cells
A:Reference number: A61180; MUID:92032882
A:Accession: A61180
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 730-838 <VIA>
R:Gonzalez, L.J.; Walker, I.D.; Sandrin, M.S.; McKenzie, I.F.C.
Immunogenetics 25, 263-266, 1987
A:Title: High sequence conservation between rat (T200) and mouse (Ly-5) leukocyte com
A:Reference number: A60933; MUID:87192931
A:Accession: A60933
A:Molecule type: protein
A:Residues: 'R', 289-298, 329, 'V', 331-336, 'Y', 'R', 364-370, 'X', 372-375, 595-608, 638-649, 6
R:Johnson, N.A.; Meyer, C.M.; Pingel, J.T.; Thomas, M.L.
J. Biol. Chem. 264, 6220-6229, 1989
A:Title: Sequence conservation in potential regulatory regions of the mouse and human
A:Reference number: A33522; MUID:89197920
A:Accession: A33522
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-22 <JOH>
A:Cross-references: GB:M22456; NID:9198755; PIDN:AAB46374.1; PID:g554185; GB:J04640;
R:Raschke, W.C.
Proc. Natl. Acad. Sci. U.S.A. 84, 161-165, 1987
A:Title: Cloned murine T200 (Ly-5) cDNA reveals multiple transcripts within B- and T-
A:Reference number: A29075; MUID:87092355
A:Accession: A29075
A:Molecule type: mRNA
A:Residues: 961-1291 <RAS>
A:Cross-references: GB:M15174; NID:g201105; PIDN:AAA40161.1; PID:g201106
R:Tung, J.
Immunogenetics 28, 271-277, 1988
A:Title: Structural features of Ly-5 glycoproteins of the mouse and counterparts in o
A:Reference number: I54450; MUID:88330145
A:Accession: I54450
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 32-73 <RES>
A:Cross-references: GB:M23241; NID:g340850; PIDN:AAA39460.1; PID:g548174
C:Genetics:
A:Gene: Ly-5
C:Superfamily: leukocyte common antigen; leukocyte common antigen cytosolic domain ho
C:Keywords: alternative splicing; glycoprotein; phosphoprotein; phosphoric monoester
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-1291/Product: protein-tyrosine-phosphatase (B-cell variant) #status predicted <M
F:24-564/Domain: extracellular #status predicted <EXT>
F:24-30,170-1291/Product: protein-tyrosine-phosphatase (T-cell variant) #status predi
F:565-586/Domain: transmembrane #status predicted <TMN>
F:587-1223/Domain: leukocyte common antigen cytosolic domain homology <LAC>
F:587-1291/Domain: intracellular #status predicted <INT>
F:664-188/Domain: protein-tyrosine-phosphatase homology <PTP>
F:664-150,160,207,211,218,253,258,290,311,322,347,416,427,457,489,520,556/Binding site
F:840/Active site: Cys (phosphocysteine intermediate) #status predicted
F:846/Binding site: substrate phosphate (Arg) #status predicted

Query Match 14.0%; Score 507; DB 1; Length 1291;
Best Local Similarity 27.8%; Pred. No. 1.8e-23;
Matches 174; Conservative 84; Mismatches 237; Indels 132; Gaps 24;

QY 1 MDQREILQKFLDEAQSCKITKEE--FANEFLKLKROSTKYKADTKYPTTVAENAKNKN 58
DB 615 MDVEPIHSDILLETYKRIADGRFLAEFQSIPIRVFSK-----FPIKDARKPHNQKN 668

Qy 282 LVYNAVLE--FKRQMDVIRDKHSGTESQAKHCIPKKNHTLQADSYGNLPK-STTKAAK 338
Db 873 LIHQALVEYNQGETEVNLSLHSCLOKLRDPSPDPLEAE--YQRLPSYRSWRTOH 930
Qy 339 MNQORTKMEIKESSDFRTSEISAKBELVHLHPAKSTSDFLNLYSFKNADT--TM 396
Db 931 IGNOEENKKRRSSNVVPYDFNRVPLKHELEM--SKESEA-----ESDESSDESDSETS 984
Qy 397 KWQTKAPFIVCEPLQKHQSLDGLSLLFEGCSNSKP---VNAAGRYFNKSVPIITTKSPF 453
Db 985 KYINASFW-----SYWPEMMIAAQ-----PIKETIGDFW 1016
Qy 454 ELIOQRETKEY-----DSKENFSYLESOPHD-----SCFVEMOAKV 490
Db 1017 QMIFQKVKVIMLTMLSGDQEVCAQWYGEKQTYGDMVYMLKDTNKSAYI-LRAFEL 1075
Qy 491 MHVSAAE-----LNYSLPYDSKHQIRNASNVK 517
Db 1076 RHSKRKEPTVYQYQCTTWKGEELPAEPKDLVTLIQNIK 1114
RESULT 11
JC6312
Protein-tyrosine-phosphatase (EC 3.1.3.48) receptor-type - human
C:Species: Homo sapiens (man)
C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 21-Jan-2000
C:Accession: JC6312
R:Yang, Y.; Gil, M.C.; Choi, E.Y.; Park, S.H.; Pyun, K.H.; Ha, H.
Gene 186, 77-82, 1997
A:Title: Molecular cloning and chromosomal localization of a human gene homologous to th
A:Reference number: JC6312; MUID:97199372
A:Accession: JC6312
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1440 <YAN>
A:Cross-references: GB:L77886
C:Superfamily: protein-tyrosine-phosphatase, receptor type mu; fibronectin type III repe
-tyrosine-phosphatase homology
C:Keywords: phosphoric monoester hydrolase
F:31-194/Domain: NAM homology <NAM>
F:209-272/Domain: immunoglobulin homology <IMM>
F:294-372/Domain: fibronectin type III repeat homology <3PR>
F:911-1131/Domain: protein-tyrosine-phosphatase homology <PTP>
Query Match 13.6%; Score 490.5; DB 2; Length 1440;
Best Local Similarity 27.08; Pred. No. 2.3e-22;
Matches 161; Conservative 91; Mismatches 222; Indels 123; Gaps 22;
Qy 21 KEAFANEFLKLRQSTKYKADTKYPTTVAENAKNIKKNRYKDILPYDSYRVELSLITSDE 80
Db 889 KEEYESFF--EGOSASW-----DVAKKDQNRKKNRYGHIIAYDHSRVILQPVEDDP 937
Qy 81 DSSYTNANPIKGYGPKAYIATQGPLSTLLDFWRMIWEYSVLIIVMACMEYEMGKKCE 140
Db 938 SSDYINANYIDGYQRPNSYIATQGPVHTVDFWRMIWEQSACIVMTNLVEVGRVKCY 997
Qy 141 RYWAEPGEMOLEFGPFSYSCAEKRKSDYIIR--TLKVKFNSETPTIYQFHYKNWPDHDV 198
Db 998 KIWPDDTEV---YGFKVCYVEMEPALBYVVRFTFLRRGYNEIRVKKQFHTGMPDHGV 1054
Qy 199 PSSDIPLEIWDVRCYQEDDSVPCTICHSAGCGRTGYCAIVDYTWMLLKDGIIPEFS 258
Db 1055 PVHATGLLSFRVAVKLSNPPSAGPIVHVCSAGAFRTGYIIVIMDMAERGVV---D 1110
Qy 259 VFSLTREMTORPSLVQTOEQVELYVNAVLEL-----FK-RQMDVIR----- 299
Db 1111 IYNCVKALRSRINNVQTEQYIFTHDAILEACLGCTAIPVCFEKAAYFMDIRDSQTN 1170
Qy 300 DKHSGTESQAKH-----CIPEKNHTLIQ---ADSYSPN--LPKSTTKAAKMN 341
Db 1171 SSKLKDEFQTLNSVTPRLQAECDSTACLP--RNHDKNRFMDMLPPDRCLPPLITIDGESSN 1229

Qy 342 QORTKM--EIKESSDFRTSEISAKBELVHLHPAKSTSDFLNLYSF----- 388
Db 1230 YINAALMDSYRQPAF-----IVTQYPLPTVKDFWRLVYDYGCTSTVMLNEV 1277
Qy 389 DKNADTTMKWQTKAF-----PIVGEPLQKHQSLDGLSLLFEGCSNSKPVNA-----A 435
Db 1278 DLSGCGPYWPEEGLRYGPIQVECMSCMDCDVINRIFRICNITRPQEGYLMVQQFYQL 1337
Qy 436 GRYNSKVPITRTTKSTPFELI-----QORETKVDKSENFYSLESQPHDSCF-----VE 484
Db 1338 GWASHREVP--GSNSFLKLLIQVEKKQOECKEGEGRTIHCINGGSRGNFCAYGIVVE 1395
Qy 485 M-QAQKVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGVYSYIPLVENPYFSS 540
Db 1396 MVKRONVV-----DVFHAVKTLRNSKPNVNEAPEQYRFCYDVALEYLES 1439
RESULT 12
S20825
protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type 6 - human
N:Alternate names: hematopoietic cell phosphatase HCP; protein-tyrosine-phosphatase 1
C:Species: Homo sapiens (man)
C:Date: 19-May-1994 #sequence_revision 08-Feb-1996 #text_change 11-Jun-1999
C:Accession: B42031; A38189; S20825; S17234; S20837
R:Yi, T.L.; Cleveland, J.L.; Ihle, J.N.
Mol. Cell. Biol. 12, 836-846, 1992
A:Title: Protein tyrosine phosphatase containing SH2 domains: characterization, prefe
A:Reference number: A42031; MUID:92123209
A:Accession: B42031
A:Molecule type: mRNA
A:Residues: 1-595 <YII>
A:Cross-references: GB:M74093
A:Experimental source: T-lymphoid cell line
A:Note: sequence extracted from GenBank
R:Plutzky, J.; Neel, B.G.; Rosenberg, R.D.
Proc. Natl. Acad. Sci. U.S.A. 89, 1123-1127, 1992
A:Title: Isolation of a src homology 2-containing tyrosine phosphatase.
A:Reference number: A38189; MUID:92141214
A:Accession: A38189
A:Molecule type: mRNA
A:Residues: 1-85, 'V', '87-595 <PLU>
A:Cross-references: GB:M77273; NID:g338079; PIDN:AAA36610.1; PID:g338080
A:Note: sequence extracted from NCBI backbone (NCBIN:79619, NCBIP:79620)
R:Shen, S.H.; Bastien, L.; Posner, B.I.; Chretien, P.
Nature 353, 868, 1991
A:Title: Corrigendum: A protein-tyrosine phosphatase with sequence similarity to the
A:Reference number: S20825
A:Accession: S20825
A:Molecule type: mRNA
A:Residues: 1, 'LSRG', 4-595 <SHE>
A:Cross-references: EMBL:X62055; NID:g35781; PIDN:CAA43982.1; PID:g35782
R:Shen, S.H.; Bastien, L.; Posner, B.I.; Chretien, P.
Nature 352, 736-739, 1991
A:Title: A protein-tyrosine phosphatase with sequence similarity to the SH2 domain of
A:Reference number: S17234; MUID:91343005
A:Accession: S17234
A:Molecule type: mRNA
A:Residues: 1, 'LSRG', 4-589, 'VPSRGSERCCPOVAMPQP' <SH2>
A:Experimental source: breast carcinoma cells
A:Note: sequence revised in reference S20805
C:Genetics:
A:Gene: GDB:PTPN6
A:Cross-references: GDB:131389; OMIM:176883
A:Map position: 12p13-12p13
C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 6; protein-tyrosine-pho
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosph
F:110-211/Domain: SH2 homology <SH2A>
F:4-98/Domain: SH2 homology <SH2B>
F:265-521/Domain: phosphatase catalytic domain #status predicted <P>
F:270-504/Domain: protein-tyrosine-phosphatase homology <PTP>
F:453/Active site: Cys (phosphocysteine intermediate) #status predicted
F:459/Binding site: substrate phosphate (Arg) #status predicted

```

Query Match      13.4%; Score 485; DB 1; Length 595;
Best Local Similarity 33.0%; Pred. No. 1.4e-22;
Matches 126; Conservative 69; Mismatches 129; Indels 62; Gaps

QY   15 QSKKITKEEFANFELKLRQSTKYKADKTYPTTVAENAKNKNRYKDILPYDSRVLS 74
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   235 ESEDTAKAGWFEBESLQKQEVK----NLHORLEGORPENKGKNRYKNILPFDHSRVLQ 290

QY   75 LITS-D-EDSSVINANFTKG-VYGP-----KAYIATOGPLSTLLDPFRMTWEYSVLIIVNA 128
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   291 GRGSIPIGSDYINANYIKNQLLGGPDENAKITYIASQCGLSATWDFMQMAQNSRIVVT 350

QY   129 CMVEYMGKKKCERYAEPGEQMLGFPGFPFSCEAEKRKSDYIRTLVKV--FNSE-TRTI 185
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   351 TREVEKGRNKCVYPWPPEVG-MQRAYGPSYNTNGEHDITTEYKLRTLOVSPLDNGDLIREI 409

QY   186 YQFHYNKPDPHDVPSSIDTLEILLWDVRCYOED--DSVPTCHCSAGCGRTGVCAIVDY 243
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   410 WHVQYLSPDHGVPSPGGVSLFDQINQRESLPHAGPIIHCSAGIGRTGTIIVI--- 466

QY   244 TWMLLKDGIIIPNFVSFL-----IREMTRPSLVOTQEQVELYVNAV---LELF 291
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   467 -----DMLMENISTKGLDCDDIDIQTKIQMVRAQRSGMVQEAQYKFYYVAIAQFIETT 519

QY   292 KROMDVIRDKHSGTESQAOKHCIEPKNHTLQADSYPNLKPSKTTAAK-----M 339
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   520 KKLELVQSQ-KGOSEYGNI-----TYPAMKNNAHAKASRTSSKHUKEDVYENL 567

QY   340 MNQORTKMEIKESSDFRTSEISAKEE 367
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   568 HTNKNREKVKKQRSADKESKSLARRK 595
    :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 13
B48148
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type gamma precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jan-2000
C:Accession: B48148
R:Barnea, G.; Siivonenoinen, O.; Shaanan, B.; Honegger, A.M.; Canoll, P.D.; D'Eustachio, M.L. Cell. Biol. 13, 1497-1506, 1993
A:Title: Identification of a carbonic anhydrase-like domain in the extracellular region
A:Reference number: A48148; MUID:93180796
A:Accession: B48148
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1442 <BAR>
A:Cross-references: GB:L09562; NID:g293773; PIDN:AAA40022.1; PID:g293774
C:Genetics:
A:Gene: Ptprg
A:Map position: 14
A>Note: extensively polymorphic
C:Superfamily: protein-tyrosine-phosphatase, receptor type gamma; carbonic anhydrase hom
ne-phosphatase homology
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; polymorphism; receptor; tran
F:60-321/Domain: carbonic anhydrase homology <CAH>
F:346-434/Domain: fibronectin type III repeat homology <3FR>
F:871-1105/Domain: protein-tyrosine-phosphatase homology <PTP>
F:1172-1396/Domain: protein-tyrosine-phosphatase homology <PTP2>
F:1057/Active site: Cys (phosphocysteine intermediate) #status Predicted
F:1063/Binding site: substrate phosphate (Arg) #status Predicted

```

Query Match      13.4%; Score 485; DB 1; Length 1442;
Best Local Similarity 25.8%; Pred. No. 5e-22;
Matches 160; Conservative 96; Mismatches 263; Indels 100; Gaps

QY 9 KFLDEAQSKKITKEEFANEFELKLKROSTKYKADKTYTTVAENAKNIKKNRKIDLPYDY 68
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 832 KHIGELYSN--SQGFSEDFEEYQRC-----ADMNTAEHSHPDNKHKNRINILAYDH 885

QY 69 SRVLESLITSDEDL--SSYINANPKGVGPKAYIATQGPLSTLTLDFFWRMIWEYSVLITV 126

```

Db	886	SRVRLRPLPKD	SKHSDYINANV	DYGNKAKAYIATQGLKSTFDFWRMWEQNTGIII	945		
Qy	127	MACMEYEMGKKCKERY	WAEPGEQLEFGFP	SYSCBAEKRSYIIRTLKV-----	176		
Db	946	MITNLVEGRKKCDY	W--PTENTEY	EGNIIVTLSTKVHACYTVRRLSVRNTKVKKQK	1003		
Qy	177	---KFNETRTTIV	QHYHKNWPDHDV	SPSSIDPILELWVRCYQOEDSDVPICHSAGCGR	233		
Db	1004	GNPKGRQNETV	QYHYHTQWPD	MGVPEYALPVLTVFRRSSAARM	DMGDPVLVHCSAGVR 1063		
Qy	234	TGVICATVDY	TWMLLDGII	PENFSYFSLIREMRTORPSL	QVOTQOYELVYNVAVLELFKR 293		
Db	1064	TGTY-IVIDS	MLQOIKD---	KSTVNVVLGFKHIRTQNV	LVQTESQYIFIHDALLE---- 1115		
Qy	294	QMDVIRDKHSGT	SOAKH-----	CIPEKNHTLQADSY	PNLPKSTTKAAKMMNQORTKM 347		
Db	1116	---AILGKET	AVSSQLHSV	NSILIPGVGGKTRLEKQFKLITCQNAKYVECF	SAQ----- 1168		
Qy	348	EIKRESSDFRT	SEI---SAKEELV	LHPAKSSTSFDFLELNSY	SFDKNADTTMKWOTKAPPI 405		
Db	1169	--KECNKEK	RNSVVAERAR	VGLAPLPGMKGDTYINASY	IMG-----YIRSNEFII 1219		
Qy	406	VGEPLQKHOS	LDLGSLLF	FECCSNKSPVNAAGRYF	NSKVPITTKSTPELTOQRETKVD 465		
Db	1220	TQHPL-PHTT	KDFRMIWDH-----	NA-----	---OIVMLPDMQSL 1252		
Qy	466	SKENFYSLE	SOPHDSCFV	EMOAKVMHVS	SAELNYSLPYDSKHQIRNASNVKHHDS	SALG 525	
Db	1253	AEDFVYWP	RES--MNCEA	FVTLSIKDR	LCLS-----NEEQII	IHDFILEA 1299	
Qy	526	VY-SYIPL	VENPYFSGW	PPSGTSSKMSLDL	PEKQDGVFPSSLLPSTSTLSFYNSHSS	584	
Db	1300	TQDDYVLEVR	HFOCPKWP	NDAPIS	STFEL---INVIKEAL	TRDGTIVHDEYGA	VS 1355
Qy	585	LSLNSPTN	ISLLNQ	PSAV	603		
Db	1356	GMICAL	FTLSQOLE	NANAV	1374		
RESULT 14							
JC5167							
protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type 11 - chicken							
N:Alternate names: phosphotyrosine phosphatase; PTP1B; PTP2c; SH-PTP2; Sy							
C:Species: Gallus gallus (chicken)							
C:Date: 21-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 16-Jul-1999							
C:Accession: JC5167							
R:Park, C.Y.; LaMontagne, K.R.; Tonks, N.K.; Hayman, M.J.							
Gene 177, 93-97, 1996							
A:Title: Cloning and expression of the chicken protein tyrosine phosphatase							
A:Reference number: JC5167; MUID:97080506							

A:Contents: erythroblast
A:Accession: JC5167
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-593 <PAR>
A:Cross-references: GB:U38620; NID:g1054939; PID:AAC60049.1; PID:g1054940
C:Comment: This enzyme plays positive roles in mitogenic signaling and early
C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 6; protein-ty
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specifi
F:6-105,112-193/Domain: SH2 #status predicted <SH2>
F:6-100/Domain: SH2 homology <SH2A>
F:112-214/Domain: SH2 homology <SH2B>
F:273-510/Domain: protein-tyrosine-phosphatase homology <PTP>
F:459/Active site: Cys (phosphotyrosine intermediate) #status predicted
F:465/Binding site: substrate phosphate (Arg) #status predicted
F:542,547,580/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match	13.4%;	Score 483.5;	DB 2;	Length 593;
Best Local Similarity	32.5%;	Pred. No. 1.7e-22;		
Matches 123;	Conservative 67;	Mismatches 119;	Indels 69;	Gaps

```
QY 14 AOSKKTKEEFANEFLLKRRQSTKYKADTKYPTTVAENAKNIKNRYKDILPYDYSRVEL 73
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 237 AETTDKVGKGFWEETLQOCEKL----LYSRKEGQROENKNKNRYKNILPFDHTRV-- 290
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 74 SLITSEDE-----SSVINANFKGVY-----GPKAVIATQGPISTLLDFWRMIWEY 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 291 --VLHDGDPNEPVSVDYINIIIMPEETKCNNSKPKSIATQGCLONTVDFWRMVFOE 348
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 SVLIIVMACMEYEMGKKKERYWAEPGEMQL-EFGPFVSCEAEKRKSDYIIRTLKV--- 176
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 349 NSRVIWMTTKEVERGSKCKVYW--PDEYSLKEYGVMVRNVKESAAHDTYTLRELKLSKV 406
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 177 -RFSNSTRITIOFHYKNWPDHVPSSIDPILLELIMDVRCYQE--DDSVPICIHCSAGCGR 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 407 GOGNTE-RIVWQYHFTWHDGVPSPDGGVLDLFLEEVHHKQESISDAGPVVVCNAGIGR 465
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 234 TGVCIAVDYTWMLLKDGIIIPENFSVFSILIREMRTQRPVSLVQEQYELVYNAV---LLEL 290
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 466 TGTFF-IVIDILIDIIREKGVDCDIDVPKTIQVRSQRSGMVQTEAQYRFIYMAVQHYIET 524
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 291 FKQMD-----VIRDKHSGTESQAKHCHICPEKNHTLOADSYSPNLPKST 333
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 525 LORRIEEOQSKRKGEYTNIKYSLSDQTSQDQSLPPCTP-----TPTCPENR 573
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 334 TRAAK-----MMNQORT 345
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 574 EDSARVYENVGLMQQKS 591
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 15

T43148

probable protein-tyrosine-phosphatase (EC 3.1.3.48) - horn shark

N:Alternate names: CD45 homolog

C:Species: Heterodontus francisci (horn shark)

C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000

C:Accession: T43148

R:Okumura, M.; Matthews, R.J.; Robb, B.; Bork, P.; Thomas, M.L.

submitted to the EMBL Data Library, August 1995

A:Reference number: Z22317

A:Accession: T43148

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1200 <OKU>

A:Cross-references: EMBL:U34750; NID:g1304393; PID:g1335805; PIDN:AAB01087.1

C:Superfamily: leukocyte common antigen; leukocyte common antigen cytosolic domain homolog
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatase

```
Query Match      13.4%; Score 483.5; DB 2; Length 1200;
Best Local Similarity 24.2%; Pred. No. 4.7e-22;
Matches 180; Conservative 116; Mismatches 307; Indels 141; Gaps 24;
```

```
QY 2 DQREIL-----QKFLDEAQAQKTKKEE-FANEFLLKRRQSTKYKADTKYPTTVAENAK 53
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 521 DEKQLLNMEPIIAEQIIDYRRKQADESLFLAEFOSIPRFSKFSVKE-----ARRGC 574
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 54 NIKKNRYKDLPLVDYSRVELSLTSDSSYINANFIKGYGPKAVIATQGPISLTLLDF 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 575 NTKNRYVDLLPYDHRVQLSPITAGEGSDYINASFIDGFNRSRKYIAAAGPKKEETSDDF 634
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 114 WRMIWEYSVLIIIVMACMEYEMGKKKERYWAEPGEMQLFEGFPFVSCEAEKRKSDYIIRT 173
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 635 WKMWVQKATIIIVMTVCEEGRPKCAQYWPMTDPSKTFGDLTVRISEQWCPDIVIRK 694
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 174 LKVKFNSE---TRTIYQHYKNWPDHVPSSIDPILLELIMDVRCYQEDDSVPICIHCSAG 230
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 695 LFTSHKSEKTPREVTHIQFIWPDHGVDPDHPHLLKLRQVNAFRNLFSGPIVHCSAG 754
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 231 CGRTGVICATVDYTWMLLKDGIIIPENFSVFSILIREMRTQRPVSLVQEQYELVYNAVLEL 290
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 755 VGRGTSGYIGIDAMMQGLEAGR-----DVIGYIVQRQRCLMVQVEAQYILIHQALLEY 810
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: May 26, 2002, 12:19:02
Job time: 5812 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2002, 12:17:55 ; Search time 39.31 Seconds
(without alignments)
681.606 Million cell updates/sec

Title: US-09-600-358A-4
Perfect score: 3615
Sequence: 1 MDQREILQFLDEAQSKIT.....RPSKSVKLRSPKSGKNFWSL 692

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2356.5	65.2	802	PTN8_MOUSE	P29352 mus musculus
2	1157	32.0	780	PTNC_HUMAN	Q05209 homo sapien
3	1130.5	31.3	775	PTNC_MOUSE	P35831 mus musculus
4	516.5	14.3	1304	CD45_HUMAN	P08575 homo sapien
5	513	14.2	1238	PTPJ_MOUSE	Q84455 mus musculus
6	508	14.1	1337	PTPJ_HUMAN	Q12913 homo sapien
7	499	13.8	1152	CD45_MOUSE	P06800 mus musculus
8	493	13.6	1255	CD45_RAT	P04157 rattus norv
9	492.5	13.6	1439	PTPK_HUMAN	Q15262 homo sapien
10	485	13.4	595	PTPG_HUMAN	P29350 homo sapien
11	485	13.4	1442	PTPG_MOUSE	Q05909 mus musculus
12	479	13.3	1452	PTPM_MOUSE	P28828 mus musculus
13	479	13.2	2316	PTP2_RAT	Q62656 rattus norv
14	478.5	13.2	1457	PTPK_MOUSE	P35822 mus musculus
15	478	13.2	2314	PTP2_HUMAN	P23471 homo sapien
16	477	13.2	1452	PTPM_HUMAN	P28827 homo sapien
17	476	13.2	1445	PTPG_HUMAN	P23470 homo sapien
18	473.5	13.1	593	PTNB_RAT	P41499 rattus norv
19	473	13.1	593	PTNB_HUMAN	Q06124 homo sapien
20	471	13.0	1912	PTPD_HUMAN	P23468 homo sapien
21	468	12.9	595	PTPM_MOUSE	P29351 mus musculus
22	467.5	12.9	1897	PTPE_HUMAN	P10586 homo sapien
23	466.5	12.9	585	PTNB_MOUSE	P35235 mus musculus
24	463.5	12.8	521	PTPL_DICDI	P34137 dictyostell
25	462.5	12.8	1301	PTPR_DROME	P35832 drosophila
26	460	12.7	802	PTPA_HUMAN	P18433 homo sapien
27	460	12.7	829	PTPA_MOUSE	P18052 mus musculus
28	456	12.6	1422	PTPG_CHICK	Q98936 gallus gall
29	452	12.5	1705	PTPO_MOUSE	P70289 mus musculus
30	446.5	12.4	1462	PTP6_DROME	P16620 drosophila
31	441	12.2	796	PTPA_RAT	Q03348 rattus norv
32	439.5	12.2	2029	LAR_DROME	P16621 drosophila
33	436.5	12.1	434	PTN1_CHICK	O13016 gallus gall

34	436	12.1	979	1	PTPN_MOUSE	Q060673 mus musculus
35	435	12.0	359	1	PTN7_RAT	P49445 rattus norv
36	434	12.0	1630	1	PTPL_DROME	P35992 drosophila
37	433	12.0	979	1	PTPN_BOVIN	P36722 bos taurus
38	433	12.0	983	1	PTPN_RAT	Q63259 rattus norv
39	432	12.0	979	1	PTPN_HUMAN	Q16849 homo sapien
40	430.5	11.9	432	1	PTN1_RAT	P20417 rattus norv
41	428	11.8	1711	1	PTPO_RAT	Q64612 rattus norv
42	427.5	11.8	435	1	PTN1_HUMAN	P18031 homo sapien
43	427.5	11.8	913	1	PTN3_HUMAN	P26045 homo sapien
44	427	11.8	926	1	PTN4_HUMAN	P29074 homo sapien
45	426.5	11.8	360	1	PTN7_HUMAN	P35236 homo sapien

ALIGNMENTS

RESULT	1
PTN8_MOUSE	
ID	PTN8_MOUSE
AC	P29352, STANDARD; PRT; 802 AA.
DT	01-DEC-1992 (Rel. 24, Created)
DT	01-DEC-1992 (Rel. 24, Last sequence update)
DT	30-MAY-2000 (Rel. 39, Last annotation update)
DE	Hematopoietic cell protein-tyrosine phosphatase 70Z-PEP (EC 3.1.3.48).
GN	PTPN8
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	MEDLINE=92236615; PubMed=1373816;
RT	Mathews R.J., Bowne D.B., Flores E., Thomas M.L.;
RT	"Characterization of hematopoietic intracellular protein tyrosine
RT	phosphatases: description of a phosphatase containing an SH2 domain
RT	and another enriched in proline-, glutamic acid-, serine-, and
RT	threonine-rich sequences";
RL	Mol. Cell. Biol. 12:2396-2405(1992).
CC	-1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC	tyrosine + phosphate.
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic.
CC	-1- TISSUE SPECIFICITY: SPLEEN, THYMUS, LYMPH NODE, AND BONE MARROW.
CC	-1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC	TYROSINE PHOSPHATASE FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; M90388; AAA39994.1; -
DR	PIR; B44390; B44390.
DR	HSSP; Q06124; 2SHP.
DR	MGD; MGI:107170; Ptpn8.
DR	InterPro; IPR000387; Tyr_phosphatase.
DR	InterPro; IPR000242; Tyr_prot_phptase.
DR	Pfam; PF00102; Y_phosphatase; 1.
DR	PRINTS; PR00700; PTPYPHPTASE.
DR	SMART; SM00194; PTPc; 1.
DR	PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR	PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR	PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW	HYDROLASE.
FT	DOMAIN 23 288 PROTEIN-TYROSINE PHOSPHATASE.
FT	ACT_SITE 227 227 BY SIMILARITY.
SQ	SEQUENCE 802 AA; 89714 MW; 0FIE45339BD4613E CRC64;
Query Match	65.2%; Score 2356.5; DB 1; Length 802;

Best Local Similarity 68.5%; Pred. No. 1.5e-137;		Matches 474; Conservative 66; Mismatches 139; Indels 13; Gaps 6;	
QY	1	MDQREILQKFLDQAQSKKITKEEFANEFLLKRSQTKYKADKTYPTTVAENAKNIKNRY	60
Db	1	MDQREILQQLKQAQKKLNSFEFLKRSQTKYKADKTYPTTVAQRPNKIKNRY	60
QY	61	KDILPYDSRVLSLTSDESDSYINANFIKGVYGPKAYIATQGPLSTLLDFWRMIWEY	120
Db	61	KDILPYDHSVLSSLTSDESDSYINAFIKGVYGPKAYIATQGPLSTLLDFWRMIWEY	120
QY	121	SVLIIVMACMEYEMGKKCRYNAEPGEMOLEFGPFSVCEAEKRKSDYIIRTLKVKFNS	180
Db	121	RILIVMACMEFEMGKKCRYNAEPGETQLQFGPFSISCEAEKRKSDYIIRTLKAFNN	180
QY	181	ETRTIYQFHYKNPDHVPSSIDPILFIWVRCYQDDSVPICIHCSACGRTGVICAI	240
Db	181	ETRIIYQFHYKNPDHVPSSIDPILQIWMRCYQDDCVPICIHCSACGRTGVICA-	239
QY	241	VDYTWMLLKDGIIPEFNSVFLSIREMTQRPVSLVQTOEQYELVYNNAVLEFKRQMDVIRD	300
Db	240	VDYTWMLLKDGIIPEFNSVFLSIQMRTQRPVSLVQTOEQYELVYSAVLEFKRQMDVIRD	299
QY	301	KHSGTESQAKHCITPEKNHTLOADSYSNLPKSTTKAAKMNQ---QRTKMEIKESSSDF	357
Db	300	NHLGREIQAQCSIPQSLTVEADSCPLDLPKNAMRDYKTTNQHSGOGAEASTGGSLGL	359
QY	358	RTSEISAKELVLHPAKSSYDFLELNYSFQDNADTKMKWTKAPPIVCEPLQKHQSLD	417
Db	360	RTSTMNAEELVLHSAKSSFNCLNCCGNKAVITRNGQARASVVGEPLOKQYSLD	419
QY	418	LGSLFLFCGCSNKPVNAGRYFNKVPITRTKSTPPELIQOQRETKEVDNSFYLESQP	477
Db	420	FGSMLFCGCSALPINTADRYHNSKGPVKRTKSTPPELIQOQRTINDLAVDGFSCLESQ	479
QY	478	HDSCFV-EMQAQKVNHYSSALNLSYDQSKHQIRNASNVKHHDSALGVYSIPLVNP	536
Db	480	HEHYSRLRELQVQVYHVSSEELNLSL-----GACDASCPVSPGALRVHLTSLAEDP	534
QY	537	YFSSMPPSGTSSKMSLDLPEKQDQGVTFPSSLLPTSSYFNSSLSLSNPTNSSL	596
Db	535	YFSSPPNSADSKMSFDLPEKQDQATSGALLPASSTSFYSNPHDSLNTLTSFSP	594
QY	597	LNQESAVLATAPRIDDEIPPLPVRTPEFIVVEAGEFSPNVPKSLSSAVKVKIGTSL	656
Db	595	LNQETAVEAPSRDDEIPPLPTEPESFIVVEAGEFSPRVTESL-PLVVFEGASPE	652
QY	657	WGTSPPKKFDDSVILRPSKSVKLRSPKSGKN	688
Db	653	CSGTSEMKSHDSVGTTPSKNVKLRSPKSDRH	683
RESULT 2			
PTNC_HUMAN STANDARD; PRT; 780 AA.			
AC	Q05209; Q16130;		
DT	01-FEB-1994 (Rel. 28, Created)		
DT	01-FEB-1994 (Rel. 28, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Protein-tyrosine phosphatase, non-receptor type 12 (EC 3.1.3.48)		
DE	(Protein-tyrosine phosphatase G1) (PTPG1).		
GN	PTPN12.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Colon;		
RX	MEDLINE=93112015; PubMed=1472029;		
RA	Takekawa M., Itoh F., Hinoda Y., Arimura Y., Toyota M., Sekiya M.,		
RA	Adachi M., Imai K., Yachi A.;		
RT	"Cloning and characterization of a human cDNA encoding a novel		
RT	putative cytoplasmic protein-tyrosine-phosphatase.";		
RL	Biochem. Biophys. Res. Commun. 189:1223-1230(1992).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=93203262; PubMed=8454633;		
RA	Yang Q.C., Tonks N.K., Sommercorn J.;		
RT	"Cloning and expression of PTP-PEST. A novel, human, nontransmembrane		
RT	protein tyrosine phosphatase.";		
RL	J. Biol. Chem. 268:6622-6628(1993).		
RN	[3]		
RP	SEQUENCE OF 59-136 FROM N.A., AND VARIANT COLON CANCER ARG-61.		
RX	MEDLINE=94156037; PubMed=7509295;		
RA	Takekawa M., Itoh F., Hinoda Y., Adachi M., Ariyama T., Inazawa J.,		
RA	Imai K., Yachi A.;		
RT	"Chromosomal localization of the protein tyrosine phosphatase G1 gene		
RT	and characterization of the aberrant transcripts in human colon		
RT	cancer cells.";		
RL	FEBS Lett. 339:222-228(1994).		
CC	-I- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein		
CC	tyrosine + phosphate.		
CC	-I- SUBCELLULAR LOCATION: Cytoplasmic.		
CC	-I- DISEASE: DEFECTS IN PTPN12 ARE FOUND IN SOME COLON CANCERS.		
CC	-I- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-		
CC	TYROSINE PHOSPHATASE FAMILY.		

CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		

CC	EMBL; D13380; BAA02648.1; -		
DR	EMBL; M93425; AAA36529.1; -		
DR	EMBL; S69184; AAB30047.2; -		
DR	PIR; JC1368; JC1368.		
DR	HSSP; Q06124; 2SHP.		
DR	MIM; 600079; -		
DR	InterPro; IPR000387; TYR_phosphatase.		
DR	InterPro; IPR000242; Tyr_prot_phptase.		
DR	Pfam; PF00102; Y_phosphatase; 1.		
DR	PRINTS; PR00700; PRTYPHPTASE.		
DR	SMART; SM00194; PTPC; 1.		
DR	PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.		
DR	PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.		
DR	PROSITE; PS00055; TYR_PHOSPHATASE_PTP; 1.		
KW	Hydrolase; Disease mutation.		
FT	DOMAIN 27 291	PROTEIN-TYROSINE PHOSPHATASE.	
FT	ACT_SITE 231 231	BY SIMILARITY.	
FT	VARIANT 61 61	K -> R (IN COLON CANCER).	
FT	/FTId=VAR_006385.		
FT	CONFLICT 121 121		
FT	V -> I (IN REF. 2).		
FT	CONFLICT 322 322		
FT	V -> I (IN REF. 2).		
FT	SEQUENCE 780 AA; 88092 MW; 48F7EE5FDAEF7512 CRC64;		
Query Match		32.0%;	Score 1157; DB 1; Length 780;
Best Local Similarity		37.0%;	Pred. No. 8e-64;
Matches 290; Conservative 112; Mismatches 237; Indels 144; Gaps 25;			
QY	1	MDQREILQKFLDQAQSKKIT-----KEEFANEFLLKRSQTKYKADKTYPTTVAENAKNIK	56
Db	1	MEQVEILRKFKIQRVQAMKSPDHNGEDNFARDFMRLRLSTKRYTEKIYPTATGKEENVK	60
QY	57	KNRYKDILPYDSRVLSLTSDESDSYINANFIKGVYGPKAYIATQGPLSTLLDFWRM	116
Db	61	KNRYKDILPFDHSRVKLTLPQSDDYIINANFIKGVYGPKAYIATQGPLTANTVIDFWRM	120
QY	117	IWEYSVLIIIVMACMEYEMGKKCRYNAEPGEMOLEFGPFSVCEAEKRKSDYIIRTLKV	176
Db	121	WVEYNVVIIVMACREFEMGKKCRYNPLYGEDPITFAPFKISCEDEQARTDYIIRTLILL	180


```
QY 177 KFNSETRTIYQHYKNWPDHVPSSIDPILELIWVRCYQEDDSVPICIHCSAGCGRTGV 236
Db 181 EFQNESRLRYQHYVWNPDPHVPSSFDSDILMSLRKYQEHEDVPICIHCSAGCGRTGA 240
QY 237 ICAIVDTYTWMLLKDGILIPENFSVSLIREMRTQPSLQVTOEQYELVYNAVLELFRQMD 296
Db 241 ICAI-DYTNWLLKAGRIPEEFNVFLIQEMRTQPSLQVTOEQYELVYNAVLELFRQMD 296
QY 297 V-----IRD--KHSGETSAQKHCIPKNNHTLQADSYSPNLPKSTTKAAKMNQORTK 346
Db 300 LYIHAQAQIADGVNEINTENMVSSIEPEKQ-----DSPPPKPPRTS-----CLVEGDAK 350
QY 347 MEIKE-----SSSDFRPTSEISAKEEL-----VLHPAKSSTSFDFLELNTS 387
Db 351 EEILQPPPEHPVPPILTPSPSAFTVTTWQNDNRYHHPKPVLMVSSSEHSAADLNRYN 410
QY 388 -----FDKNADTTMKQTKAFPIVGEPLQKHOSLDGLSLLFEG----- 425
Db 411 KSTELPGKNESTIEQDKLERLMSFEIKKVPLOQEP-----KSFQDNTLLNRCGHAIKIS 466
QY 426 ---C---SNSKP-----VNAAGRYNSKVPITRTKS-----TPFELIQORETK----- 462
Db 467 ASPIADKISKPOELSSDNLVGTDSNCSVDCSVTOSKNVSVTPPEESONSDFPPRDLR 526
QY 463 EVDKSNFYLESQPSHDSFCFVEMQAOKVMHVSSAEALNY-----SLPYDSKHKQIRNASNV 516
Db 527 PLDEKGHVWTFHGPENA---IPIDLSEGNSSDINYNQTRKTVSLVTPSPTTQVETPDLV 582
QY 517 KHHDSALGVYIPIVLENYPSSWPGTSSKMSLDLPEK-ODGVYFVSSLLPTSTSL 575
Db 583 DHNTS-----PLFRTPLSFTNPLHSDDS-----DSDERNSDGAVTQKNTNISTASAT 630
QY 576 FSYNSHSSLNPSNTNISLLNOESAVLATAPR-----IDDETPPPLPVRTPESTFVVEE 631
Db 631 VSAATSTESISIRKVLPMISARINIACTHSGAEKDVVDSEDSPPLPERTPESTFVASE 690
QY 632 AGEFSPNVKSLSSAVKVKIGTSLWEG-----GTSEPKFDDSVILRPKSKVLRSKPSG 686
Db 691 -----HNTPE-----VRSESELQSPERSQKKSEG---LITSENEKCDHPAGS 730
QY 687 KNF 689
Db 731 IHY 733

RESULT 3
PTNC_MOUSE STANDARD; PRT; 775 AA.
AC P35831;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Protein-tyrosine phosphatase P19 (EC 3.1.3.48) (P19-PTP) (MPTP-PEST).
GN PTPN12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92272714; PubMed=1590786;
RA den Hertog J., Pals C.E., Jonk L.J., Kruijer W.;
RT "Differential expression of a novel murine non-receptor protein
RT tyrosine phosphatase during differentiation of P19 embryonal
RT carcinoma cells.";
RL Biochem. Biophys. Res. Commun. 184:1241-1249(1992).
RN [2]
RP REVISIONS TO 297-416.
RX MEDLINE=93112015; PubMed=1472029;
RA Takekawa M., Itoh F., Hinoda Y., Arimura Y., Toyota M., Sekiya M.,
RA Adachi M., Imai K., Yachi A.;
RT "Cloning and characterization of a human cDNA encoding a novel
RT putative cytoplasmic protein-tyrosine-phosphatase.";
```

```
RL Biochem. Biophys. Res. Commun. 189:1223-1230(1992).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE=95289971; PubMed=7772023;
RA Charest A., Wagner J., Shen S.H., Tremblay M.L.;
RT "Murine protein tyrosine phosphatase-PEST, a stable cytosolic protein
RT tyrosine phosphatase.";
RL Biochem. J. 308:425-432(1995).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O -> protein
CC tyrosine + phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X63440; CAA45037.1; ALT_SEQ.
DR EMBL; X86781; CAA60477.1; -.
DR HSP; O06124; 2SHP.
DR MGD; MGI:104673; Ptpn12.
DR InterPro; IPR000387; TYR_phosphatase.
DR InterPro; IPR000242; Tyr_prot_phphatase.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00700; PRTYPHPHTASE.
DR SMART; SM00194; PTPc; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase.
FT DOMAIN 27 291 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 231 231 BY SIMILARITY.
FT CONFLICT 296 296 K -> N (IN REF. 1).
FT CONFLICT 328 332 KODSP -> DETS (IN REF. 1).
FT CONFLICT 380 380 W -> V (IN REF. 1).
SQ SEQUENCE 775 AA; 86992 MW; 7106D73F5014E411 CRC64;

Query Match 31.3%; Score 1130.5; DB 1; Length 775;
Best Local Similarity 36.6%; Pred. No. 3.4e-62;
Matches 280; Conservative 110; Mismatches 165; Indels 209; Gaps 28;

QY 1 MDQREILQKFLDEAQSCKIT----KEEFANEFLLKRSQSTKYKADKTYPTTVAENAKNIK 56
Db 1 MEQVEILRRFIQORVQAMKSPDHNGEDNFARDFMRLRLSTKYRTEKIYPTATGEKENVK 60
QY 57 KNRYKDIPLDYRSVELSLITSDSDSYINANPIKGVGPKAYIATQGPLSTLLDFWRM 116
Db 61 KNRYKDIPLDFHSRVKLTSLKTPSQSDSYINANPIKGVGPKAYVATQGPFRNTVIDFWRM 120
QY 117 IWEYSVLIVMACMEYEMGKKCYWAEFGEMQLFEGPFSYSCAEKKRSDYIIITLKV 176
Db 121 IWEYNVMIVMACREFEMGKKCYWPLYGDEPTIFAFPKISCENEQARTFIITLL 180
QY 177 KFNSETRTIYQHYKNWPDHVPSSIDPILELIWVRCYQEDDSVPICIHCSAGCGRTGV 236
Db 181 EFQNESRLRYQHYVWNPDPHVPSSFDSDILMSLRKYQEHEDVPICIHCSAGCGRTGA 240
QY 237 ICAIVDTYTWMLLKDGILIPENFSVSLIREMRTQPSLQVTOEQYELVYNAVLELFRQMD 296
Db 241 ICAI-DYTNWLLKAGRIPEEFNVFLIQEMRTQPSLQVTOEQYELVYNAVLELFRQMD 299
QY 297 V-----IRDKHSGETSAQKHCIPKNNHTLQADSYSPNLPKSTTKAAKMNQORTK 348
Db 300 LYIHAQAQIADGVNEINTENMVSSIDSE-----KQDPPPKPPRTS-----CLVEGDAKEE 351
QY 349 IKE-----SSSDFRPTSEISAKEEL-----EISAKEELVLPKAKSSTSF 379
```

```
Db 352 ILQPPHPVPPILTPSPSPSAFTVTWQSDRYHFKPVLUHMASPEQ---HPA----- 402
Qy 380 DFLELNYSEFDKNADTMKW-----QTKAFPIVGRPELOK-HOSLDIGSLIF 423
Db 403 ---DLNRSYDASD---QWKGSESAIEHIDKKLERNUSFEKKVPLQEGPKSFDGNTLLN 456
Qy 424 EGCNSKVPVNAAGRYFNKSVPIRTKTKSPFEL-----IQOREYKEV----- 464
Db 457 RG--HAIKISA-----SSSVYDRT-SKPQELSAGALKAVDDVSQNSCADCSAAHSHRAAE 508
Qy 465 DSKENFVLESQPHSDSCFVENQAKVMHVSAAELNYSI-----PYDSK 507
Db 509 SSESQSNSHTPPRPDC---LPLDKKGHTV---WSLHGPNATVPDSDPGKSPDNHS 560
Qy 508 HQIRNASN-----VKHDDSSALGVSYIPLAVENP-----YFSSWPPPGCTS 547
Db 561 QTLKTVSTPNSTABEEAHNDTEHNS-----PULKAPLFTNPLHSDDDHWSGGGS 612
Qy 548 SKMSLDLPEKQDGVFPSSLLPTSTSLFSYNSHSLSLNSPTNMISSLLNQESAV----- 603
Db 613 S-----DGAV-----TRNKTSISTASAT-VSPASSAESACHERV 645
Qy 604 -----LATAPR-----IDETPPPLPVRTPEFIVE 630
Db 646 LPMSTARQEVAGTGHSGAEKDADVSESPPLPERTPEFSLAD 689

RESULT 4
CD45_HUMAN STANDARD; PRT; 1304 AA.
ID CD45_HUMAN
AC P08575;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Leukocyte common antigen precursor (BC 3.1.3.48) (L-CA) (CD45 antigen)
DE (T200).
GN PTPRC OR CD45.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymphocytes;
RX MEDLINE=88061067; PubMed=2824653;
RA Streuli M., Hall L.R., Saga Y., Schlossman S.F., Saito H.;
RT "Differential usage of three exons generates at least five different
RT mRNAs encoding human leukocyte common antigens.";
RL J. Exp. Med. 166:1548-1566(1987).
RN [2]
RP FUNCTION.
RX MEDLINE=89017162; PubMed=2845400;
RA Charbonneau H., Tonks N.K., Walsh K.A., Fischer E.H.;
RT "The leukocyte common antigen (CD45): a putative receptor-linked
RT protein tyrosine phosphatase.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:7182-7186(1988).
RN [3]
RP MUTAGENESIS.
RX MEDLINE=90316093; PubMed=1695146;
RA Streuli M., Krueger N.X., Thai T., Tang M., Saito H.;
RT "Distinct functional roles of the two intracellular phosphatase like
RT domains of the receptor-linked protein tyrosine phosphatases LCA and
RT LAR.";
RL EMBO J. 9:2399-2407(1990).
CC -!- FUNCTION: REQUIRED FOR T-CELLS ACTIVATION THROUGH THE ANTIGEN
CC RECEPTOR. THE FIRST PPASE DOMAIN HAS ENZYMAIC ACTIVITY, WHILE
CC THE SECOND ONE SEEMS TO AFFECT THE SUBSTRATE SPECIFICITY OF THE
CC FIRST ONE.
CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS: AT LEAST 8 ISOFORMS ARE PRODUCED BY
```

Db 666 PIKARPPFNQNKRYVDILPYDYNRVLSEINGDAGSNYNASYIDGFKPRKYIAAQ 725
QY 105 PLSTLLDFWRMIWEYSVLIIIVMACMEYEMGKKCEYWAEPGEQMLEFPGFVSCEAEK 164
Db 726 PRDVTDDFRMIWEQATVIVWTRCEENRNKCAEYFWSMEGTRAFGDVVVKINQHK 785
QY 165 RKSDYIIRTLKV---KFNSETRTIYQPHYKNWPDHVPSSIDPILELIWVRCYQEDDSV 221
Db 786 RCPDYIIQKLVNKKKATGREVTHIQFTSWPDHGVPEPHLLKLLRRVNAFNEFSFG 845
QY 222 PICHCAGCGRGVCAIVDYTWMLKDGIIIPEN-FSVFSLIREMTQRPVSLVQTOEQY 280
Db 846 PIVVHCASAGVGTGTIGI-----DAMLEGLEAENKVDYVGYVVKLRQRCLMVQVEAQY 900
QY 281 ELVYNVAVLEL--FKRQMDVIRDKHSGTESOAKHCIPKNTKLOADSYSPNLPK-STTKAA 337
Db 901 ILIHAQVLEVNQGETEVNSELPHYLHNMKKRDPPEPSPLEAE--FQRLPSYRSWRTQ 958
QY 338 KMNQORTMEIKESSDFRTSEISAKELVL-----HPAKSSTSDFFLELNYSFDKNA 392
Db 959 HIGNOEENKSNRNSVPIPDYNRVPLKHELENSKEHSDSDSDSDSDSEPSKYINA 1018
QY 393 DTMK-HQTKAPPIVGEPLQKHSLODGLSLF-----EGCSNSKPVN 433
Db 1019 SFIMSYWKPVEVMAAAGPL-KETIGDFWQMFQKVKVIVMLTELKHGDQIECAQ---YW 1074
QY 434 AAGR--YFNSKVPITRT-KSTPEL--IQORETKVDNSKENFSY-----LESOPHD 479
Db 1075 GEGQTYGDIQVLEKDFKSTTTLRVFELRHSKRKDSRTVYQYQYNNWSVEQLPAEPKE 1134
QY 480 -SCFVEMQAQVHVHSAELNYSILPYDSKHKQIRNASNVKHHDSALGVY 527
Db 1135 LISMIQVVKQKPKNSSEGN-----KHKSTPLLIH 1166

RESULT 5
PTPJ_MOUSE
ID PTPJ_MOUSE STANDARD; PRT; 1238 AA.
AC Q64455;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE protein-tyrosine phosphatase eta precursor (EC 3.1.3.48) (R-PTP-eta)
DE (HPTP beta-like tyrosine phosphatase).
GN PTPRJ OR BYP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MRL-LPR/LPR;
RX MEDLINE=96140699; PubMed=8549806;
RA Kuramochi S., Matsuda S., Matsuda Y., Saitoh T., Ohsugi M.,
Yamamoto T.;
RT "Molecular cloning and characterization of ByP, a murine
RT receptor-type tyrosine phosphatase similar to human DEP-1";
RT FEBS Lett. 378:7-14(1996).
CC -|- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -|- SUBCELLULAR LOCATION: Type I membrane protein.
CC -|- TISSUE SPECIFICITY: EXPRESSED IN EVERY TISSUE EXAMINED.
CC -|- SIMILARITY: CONTAINS 1 PROTEIN-TYROSINE PHOSPHATASE DOMAIN.
CC -|- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; D45212; BAA08146.1; -.
DR HSP; P18052; 1YFO.
DR GSD; MGI:104574; Ptprj.
DR InterPro; IPR003961; Ptprj.
DR InterPro; IPR000387; Tyr_phosphatase.
DR InterPro; IPR000242; Tyr_prot_phptase.
DR Pfam; PF00041; fn3; 6.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00700; PRTYPPHPTASE.
DR SMART; SM00060; FN3; 5.
DR SMART; SM00194; PTPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00055; TYR_PHOSPHATASE_PTP; 1.
KW Signal; Glycoprotein; Transmembrane; Repeat; Hydrolase.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 1238 PROTEIN-TYROSINE PHOSPHATASE ETA.
FT DOMAIN 29 876 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 877 897 POTENTIAL.
FT DOMAIN 898 1238 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 40 112 FIBRONECTIN TYPE-III 1.
FT DOMAIN 121 260 FIBRONECTIN TYPE-III 2.
FT DOMAIN 268 348 FIBRONECTIN TYPE-III 3.
FT DOMAIN 356 434 FIBRONECTIN TYPE-III 4.
FT DOMAIN 442 518 FIBRONECTIN TYPE-III 5.
FT DOMAIN 529 608 FIBRONECTIN TYPE-III 6.
FT DOMAIN 966 1238 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 1140 1140 BY SIMILARITY.
FT CARBOHYD 62 62 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 78 78 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 85 85 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 90 90 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 110 110 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 114 114 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 164 164 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 173 173 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 182 182 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 207 207 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 244 244 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 253 253 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 278 278 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 317 317 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 366 366 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 379 379 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 452 452 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 488 488 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 506 506 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 538 538 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 572 572 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 576 576 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 662 662 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 668 668 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 685 685 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 691 691 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 725 725 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 811 811 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 838 838 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1238 AA; 136782 MW; 939479EDC8016835 CRC64;

Query Match 14.2%; Score 513; DB 1; Length 1238;
Best Local Similarity 39.0%; Pred. No. 5.9e-24;
Matches 115; Conservative 50; Mismatches 110; Indels 20; Gaps 7;

QY 15 QSKKITKEEFANFLKLKROST-----KYADK---TYPTVAENAKNIKNRYDILP 65

Qy 126 VMACMEYEMGKKCERYWAEPCEMOLEFSPVSCAEAKRSDYIIRTLKVK--FNSETR 183
 Db 1136 IMLTKVCEQGRKCEYW--FSKQADYGDITVAMTSEIVLPETWIRDTFTVNIQTSESH 1193
 Qy 184 TIYQFHYKNWDPHDVPSSIDILELWDRY--QEDSVFICHCAGCGRTGVICAV 241
 Db 1194 PLRQPHFTSWPDHGVDPDLDLLINFRYLVRDYMKQSPPELVRHCSAGVGRGTFTIAID 1253
 Qy 242 DYTMLLKGDIIPENFVSFLIREMRTORPSIVTOEQYELVYNVAVLELFRQMD 296
 Db 1254 RLIIQIENENTV----DVIGIVIDLURHRLMVLQVOTEDQYVFLNQCVDIVRSQKD 1304

RESULT 7
 CD45_MOUSE
 ID CD45_MOUSE STANDARD; PRT; 1152 AA.
 AC P06800;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Leukocyte common antigen precursor (EC 3.1.3.48) (L-CA) (Lymphocyte
 common antigen Ly-5) (CD45) (T200).
 GN PTPRC OR Ly-5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8631368; PubMed=2944116;
 RA Saga Y., Tung J.-S., Shen F.-W., Boyse E.A.;
 RA "Sequences of Ly-5 cDNA: isoform-related diversity of Ly-5 mRNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:6940-6944(1986).
 RN [2]
 RP REVISIONS.
 RA Saga Y., Tung J.-S., Shen F.-W., Boyse E.A.;
 RL Proc. Natl. Acad. Sci. U.S.A. 84:1991-1991(1987).
 RN [3]
 RP SEQUENCE OF 10-124 FROM N.A.
 RC TISSUE=T-cell;
 RX MEDLINE=86042665; PubMed=3864163;
 RA Shen F.-W., Saga Y., Litman G., Freeman G., Tung J.-S., Cantor H.,
 RA Boyse E.A.;
 RT "Cloning of Ly-5 cDNA";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:7360-7363(1985).
 RN [4]
 RP SEQUENCE OF 822-1152 FROM N.A.
 RX MEDLINE=87092355; PubMed=2948186;
 RA Raschke W.C.;
 RT "Cloned murine T200 (Ly-5) cDNA reveals multiple transcripts within
 B- and T-lymphocyte lineages";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:161-165(1987).
 CC -1- FUNCTION: REQUIRED FOR T-CELLS ACTIVATION THROUGH THE ANTIGEN
 RECEPTOR. THE FIRST PTPASE DOMAIN HAS ENZYMIC ACTIVITY, WHILE
 THE SECOND ONE SEEMS TO AFFECT THE SUBSTRATE SPECIFICITY OF THE
 FIRST ONE.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O -> protein
 tyrosine + phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: A NUMBER OF ISOFORMS ARE PRODUCED BY
 ALTERNATIVE SPLICING.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION IS RESTRICTED TO THE HEMATOPOIETIC
 COMPARTMENT OF DEVELOPMENT.
 CC -1- PTM: HEAVILY N- AND O-GLYCOSYLATED.
 CC -1- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>)

CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL; M14342; AAA39458.1; -
 DR EMBL; M1934; AAA39461.1; -
 DR EMBL; M15174; AAA40161.1; -
 DR PIR; A29381; A29381.
 DR HSP; P18052; 1YFO.
 DR MGD; MGI:97810; Ptprc.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR000387; Tyr_phosphatase.
 DR InterPro; IPR000242; Tyr_prot_phptase.
 DR Pfam; PF00041; fn3; 3.
 DR PRINTS; PR00102; Y_phosphatase; 2.
 DR SMART; SM00194; PTPC; 2.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 2.
 DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 2.
 DR PROSITE; PS00055; TYR_PHOSPHATASE_PTP; 2.
 KW Glycoprotein; Transmembrane; Phosphorylation; B-cell; T-cell; Repeat;
 KW Alternative splicing; Hydrolase; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 1152 LEUCOCYTE COMMON ANTIGEN.
 FT DOMAIN 24 425 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 426 447 POTENTIAL.
 FT DOMAIN 448 1152 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 233 329 FIBRONECTIN TYPE-III 1.
 FT DOMAIN 330 421 FIBRONECTIN TYPE-III 2.
 FT DOMAIN 520 769 PROTEIN-TYROSINE PHOSPHATASE 1.
 FT DOMAIN 811 1084 PROTEIN-TYROSINE PHOSPHATASE 2.
 FT ACT_SITE 701 701 BY SIMILARITY.
 FT ACT_SITE 1016 1016 BY SIMILARITY.
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 79 79 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 114 114 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 119 119 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 288 288 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 318 318 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 379 379 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1152 AA; B4D956B4E32EA812 CRC64;
 Query Match 13.8%; Score 499; DB 1; Length 1152;
 Best Local Similarity 27.6%; Pred. No. 3.8e-23;
 Matches 173; Conservative 83; Mismatches 239; Indels 132; Gaps 24;
 Qy 1 MDQREILQKFLDEAQSCKITKEE--FANEFLKLRQSTKYKADKTYPTTVAENAKTKKN 58
 Db 476 MDVEPIHSDILETVYKRIADGRGLAEFQSIPIRVFSK-----FPIKDARKPHNQKN 529
 Qy 59 RYKILPYDYSRVELSLTSDSDSYINANFIKGYGPKAYIATQGPLSTLLDFWRMIW 118
 Db 530 RYVDILPYDYNRVELSEINGDAGSTYINASYIDGFKPERKYIAAAGPRDETVDFWRMIW 589
 Qy 119 EYSLVLIIVMACMEVEMGKKCKERYWAEPCEMOLEFSPVSCAEAKRSDYIIRTLKVKF 178
 Db 590 EQKATVIVWVTRCEGGRNCKAEYWPSEEGTFAKDIIVTINDHKRCPDYIIOKLVAVH 649
 Qy 179 NSET---RTIYQFHYKNWDPHDVPSSIDILELWDRYQEDSVFICHCAGCGRTG 235
 Db 650 KKERATGREVTHIQTSWPDHGVDPEDPHLLKLRNRVNAFNSFFSGPIVWHCSAGVGRGT 709
 Qy 236 VICAIYDVTWMLLKGDIPE-NFSVFLIREMRTORPSIVTOEQYELVYNVAVLEL--FK 292
 Db 710 TYIGI-----DAMLEGLEAEGKVDVGYVVKLRQRCILMQVOVEAQYILIHQALVEYNQFG 764
 Qy 293 RQMDVIRDKHSGTSGQAKHCIPKHNHTLOADSYSPNLPK-STTKAAMNMNQRTKMEIKE 351

```
Db 765 EQVNLSELHSLHNMKRDPPSPDPSPREAE--YQRLPSYRSWRTQHTGNCQENKKRN 822
QY 352 SSSEDFRSEISAKBELVL---HPAKSTSPDFLELNSYDPKNDATMTMKWOTKAPFVIG 407
Db 823 SNVYPDNRVPLKHELEMSKESEPESSDD-----DSDSEETSKYINASFVM-- 872
QY 408 EPLQKHQSILGSLLEFGCSNSKP---VNAAGRYFNKSVPIITRTKSTPFELIQQRREKV 464
Db 873 -----SYNKPWMTAAQ-----PLKETIGDFWQMFQKRVKI 906
QY 465 -----DSKENFSYLESQPHDSCFVEMQAQKVMHVSSAELNYSL-PY 504
Db 907 VMLTELNGDQVCAQYWGEGKQTYGDME-----VEM-----KDTNRASA---YTLRTF 952
QY 505 DSKHOIRNASNVKHHDSALGALVYSIPLVNPYFSSWP-----PSTGSSKMSL--DLPEKQ 558
Db 953 EURLSKRPRT-----VYQY-----QCTTWKGEELPAEPKDLVSMIQDLKOK- 995
QY 559 DGTVPFSSLLPTSSLSFSYNSHSSL 585
Db 996 -----LPKASPEGMKYH-KHASI 1012

RESULT 8
CD45_RAT STANDARD; PRT: 1255 AA.
AC P04157;
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Leukocyte common antigen variant 4 precursor (BC 3.1.3.48) (L-CA)
DE (CD45) (T200) (Fragment).
GN PTPRC.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Barclay A.N., Jackson D.I., Willis A.C., Williams A.F.;
RL Submitted (MAY-1987) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 190-1255 FROM N.A.
RX MEDLINE=85201691; PubMed=3158393;
RA Thomas M.L., Barclay A.N., Gagnon J., Williams A.F.;
RT "Evidence from cDNA clones that the rat leukocyte-common antigen
RT (T200) spans the lipid bilayer and contains a cytoplasmic domain of
RT 80,000 Mr.";
RL Cell 41:83-93(1985).
RN [3]
RP ALTERNATIVE SPLICING.
RX MEDLINE=87275817; PubMed=2440674;
RA Barclay A.N., Jackson D.I., Willis A.C., Williams A.F.;
RT "Lymphocyte specific heterogeneity in the rat leukocyte common
RT antigen (T200) is due to differences in polypeptide sequences near
RT the NH2-terminus.";
RL EMBO J. 6:1259-1264(1987).
CC -!- FUNCTION: REQUIRED FOR T-CELLS ACTIVATION THROUGH THE ANTIGEN
CC RECEPTOR. THE FIRST PTASE DOMAIN HAS ENZYMIC ACTIVITY, WHILE
CC THE SECOND ONE SEEMS TO AFFECT THE SUBSTRATE SPECIFICITY OF THE
CC FIRST ONE.
CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS: THERE ARE DIFFERENT VARIANTS OF L-CA, WHICH
CC ARISE BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: VARIANTS 4 AND 3 ARE FOUND IN THE LYMPH NODE,
CC VARIANTS 1 AND 2 ARE FOUND IN THYMOCYTE AND LYMPH NODE.
CC -!- PTM: HEAVY N- AND O-GLYCOSYLATED.
CC -!- PTM: THE CYTOPLASMIC DOMAIN CONTAINS POTENTIAL PHOSPHORYLATION
CC SITES.
CC -!- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.

-!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; Y00065; CAA68272.1; -
EMBL; Y00065; CAA68273.1; -
EMBL; Y00065; CAA68274.1; -
EMBL; Y00065; CAA68275.1; -
EMBL; M25820; AAA41518.1; -
EMBL; M25821; AAA41519.1; -
EMBL; M25822; AAA41520.1; -
EMBL; M25823; AAA41521.1; -
PIR; A60241; TDRTLT.
HSSP; P18052; 1VFO.
InterPro; IPR003961; FN_III.
InterPro; IPR000387; TYR_phosphatase.
InterPro; IPR000242; Tyr_prot_phptase.
Pfam; PF00041; fn3; 2.
SMART; SM00060; FN3; 2.
SMART; SM00194; PTPC; 2.
PROSITE; PS00383; TYR_PHOSPHATASE_1; 2.
PROSITE; PS00056; TYR_PHOSPHATASE_2; 2.
PROSITE; PS00055; TYR_PHOSPHATASE_PTP; 2.
KW Glycoprotein; Transmembrane; Phosphorylation; B-cell; T-cell; Repeat;
KW Alternative splicing; Hydrolase; Signal.
FT NON_TER 1 1
FT SIGNAL 1 5
FT CHAIN 6 1255
FT DOMAIN 6 528
FT TRANSMEM 529 550
FT DOMAIN 551 1255
FT DOMAIN 341 432
FT DOMAIN 433 524
FT DOMAIN 623 872
FT DOMAIN 914 1187
FT ACT_SITE 804 804
FT ACT_SITE 1119 1119
FT CARBOHYD 44 44
FT CARBOHYD 124 124
FT CARBOHYD 135 135
FT CARBOHYD 146 146
FT CARBOHYD 160 160
FT CARBOHYD 182 182
FT CARBOHYD 227 227
FT CARBOHYD 232 232
FT CARBOHYD 253 253
FT CARBOHYD 264 264
FT CARBOHYD 309 309
FT CARBOHYD 315 315
FT CARBOHYD 353 353
FT CARBOHYD 356 356
FT CARBOHYD 453 453
FT CARBOHYD 484 484
FT VARSPPLIC 12 53
FT VARSPPLIC 53 143
FT VARSPPLIC 103 143
FT CONFLICT 38 38
SQ SEQUENCE 1255 AA; 141208 MW; C257CBD2A355BCEA CRC64;

LEUKOCYTE COMMON ANTIGEN VARIANT 4.
EXTRACELLULAR (POTENTIAL).
POTENTIAL.
CYTOPLASMIC (POTENTIAL).
FIBRONECTIN TYPE-III 1.
FIBRONECTIN TYPE-III 2.
PROTEIN-TYROSINE PHOSPHATASE 1.
PROTEIN-TYROSINE PHOSPHATASE 2.
BY SIMILARITY.
BY SIMILARITY.
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
MISSING (IN ISOFORM 3).
MISSING (IN ISOFORM 2).
MISSING (IN ISOFORM 1).
MISSING (IN ISOFORM 3).
S -> R (IN REF. 3).

Query Match 13.6%; Score 493; DB 1; Length 1255;
Best Local Similarity 27.3%; Pred. No. 1e-22;
Matches 158; Conservative 89; Mismatches 22; Indels 110; Gaps 18;

QY 2 DOREILQK-----FLDEAQSKKITKEE--FANEFLKQRQSTKYKADTKY 44
```

```
Db 565 EQELVERDEEKQLINVDPIHSDLLLETKYKRIADCEGRLEAFQSPRVFSK-----F 618
QY 45 PTTVAENAKIKNRKYDILPYDSRVELSLIYSDDESSYNANFNKGVGPKAYATQ 104
Db 619 PDKARKSQNRKRVYDILPYDYNRVELSEINGDAGSTYINASYIDGFEPRKYIAAQ 678
QY 105 PLSTTLDFWRMTWEXSVLIIVMACMEYEMGKKCKERYWAEPCGEMQLEFGPFVSCEAEK 164
Db 679 PRDETVDFFWKMLWEQKATVIMVTRCEEGRNRKCAEYWPCEEGTTFPRDVTVTINDHK 738
QY 165 RKSDYIIRTLKVFNSSET---RTIYOFHYKNWPDHVPSSIDPILLELWDRVCYQEDDSV 221
Db 739 RCPDYIIQKLSIAHKKEKATGREVTHIQFTSWPDHGVPDPHLLKLRVRNFAFSNFFSG 798
QY 222 PICIHSGAGGRGVCAIVDYTWMLLKDGIIPEFNSVFSFLIREMTQRSLIVQTQOYE 281
Db 799 PIYVHCSAGVGRGTIGYIDAMLESAEGKV---DVYGVVNLRRQRCLMVQVQAYI 854
QY 282 LVYNAVLEL--FKRQMDVIRDKHSGTESQAKHCIPKHNHTLOADSYSPNLPK-STTKAAK 338
Db 855 LIHQALVEYNQFGETEVNLSHSCQLNKKRDPSPDPSPLEAE--YQRLPSYRSWRTOH 912
QY 339 MMNQRTKMEIKBSSSFDFTSISAKEELVLPKAKSSTSFDELELNYSPDKNADT--TM 396
Db 913 IGNOEENKKNRSSNVYPYDNVRPLKHELM--SKESEA---ESDESDSDSEETS 966
QY 397 KWOTKAFPIVGEPLQKHQSLDLGSLLEFGCSNKP---VNAAGRYNSKVPITRTKSTPF 453
Db 967 KYINASFVM-----SYWKPEMMIAAQQ-----PLKETIGDFW 998
QY 454 ELIQORETKEV-----DSKENFSYLESQPHD---SCFVEMQAQKV 490
Db 999 OMIFQKRVKVVIMLTMSGDQEVCAQYWGEGQKTGDMVEVMLKDKNTKSAYI-LRAFEL 1057
QY 491 MHYSSAE-----LNYSILPYDSKHQIRNASNVK 517
Db 1058 RSHKREKPRTVYQCTTWKGEELPAEPKDLVLTIQNIK 1096

RESULT 9
PTPK_HUMAN
ID PTPK_HUMAN STANDARD; PRT: 1439 AA.
AC Q15262; Q14763;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein-tyrosine phosphatase kappa precursor (EC 3.1.3.48) (R-PTP-
DE kappa).
GN PTPRK OR PTPK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96279245; PubMed=8663237;
RA Fuchs M., Mueller T., Lerch M., Ullrich A.;
RT "Association of human protein-tyrosine phosphatase kappa with members
RT of the armadillo family.";
RL J. Biol. Chem. 271:16712-16719(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Foreskin;
RX MEDLINE=97199372; PubMed=9047348;
RA Yang Y., Gil M.C., Choi E.Y., Park S.H., Pyun K.H., Ha H.;
RT "Molecular cloning and chromosomal localization of a human gene
RT homologous to the murine R-PTP-kappa, a receptor-type protein
RT tyrosine phosphatase.";
RL Gene 186:77-82(1997).
CC -1- FUNCTION: REGULATION OF PROCESSES INVOLVING CELL CONTACT AND
CC ADHESION SUCH AS GROWTH CONTROL, TUMOR INVASION, AND METASTASIS.
CC FORMS COMPLEXES WITH BETA-CATENIN AND GAMMA-CATENIN/PLAKOGLOBIN.
CC
```

```
CC BETA-CATENIN MAY BE A SUBSTRATE FOR THE CATALYTIC ACTIVITY OF PTP-
CC KAPPA.
CC CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O -> protein
CC tyrosine + phosphate.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN; AT ADHERENS
CC JUNCTIONS.
CC -1- TISSUE SPECIFICITY: HIGH LEVELS IN LUNG, BRAIN AND COLON; LESS IN
CC LIVER, PANCREAS, STOMACH, KIDNEY, PLACENTA AND MAMMARY CARCINOMA.
CC -1- PTM: THIS PROTEIN UNDERGOES PROTEOLYTIC PROCESSING.
CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.
CC -1- SIMILARITY: CONTAINS 4 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Z70660; CAA94519.1; -.
CC EMBL: L77886; AAC37599.1; -.
CC HSSP: P28827; 1RPM.
CC MTM: 602545; -.
CC InterPro: IPR003961; FN_III.
CC InterPro: IPR003962; FNIII_repeat.
CC HSSP: P28827; 1RPM.
CC InterPro: IPR003599; Ig.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR000998; MAM.
CC InterPro: IPR000387; TYR_phosphatase.
CC InterPro: IPR000242; Tyr_prot_phphatase.
CC Pfam: PF00041; fn3; 2.
CC Pfam: PF00047; ig; 1.
CC Pfam: PF00629; MAM; 1.
CC PRINTS: PR00102; Y_phosphatase; 2.
CC PRINTS: PR00014; ENTPEI11.
CC PRINTS: PR00020; MAMDOMAIN.
CC SMART: SM00700; PRTYPHPTASE.
CC SMART: SM00060; FN3; 2.
CC SMART: SM00409; IG; 1.
CC SMART: SM00137; MAM; 1.
CC SMART: SM00194; PTPC; 2.
CC PROSITE: PS00383; TYR_PHOSPHATASE_1; 2.
CC PROSITE: PS50056; TYR_PHOSPHATASE_2; 2.
CC PROSITE: PS50055; TYR_PHOSPHATASE_PTP; 2.
CC PROSITE: PS00740; MAM_1; 1.
CC PROSITE: PS50060; MAM_2; 1.
CC Hydrolase; Receptor; Glycoprotein; Signal; Transmembrane;
KW Immunoglobulin domain; Repeat.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 1439 PROTEIN-TYROSINE PHOSPHATASE KAPPA.
FT DOMAIN 27 752 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 753 774 POTENTIAL.
FT DOMAIN 775 1439 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 31 194 MAM.
FT DOMAIN 209 277 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 297 378 FIBRONECTIN TYPE-III 1.
FT DOMAIN 393 474 FIBRONECTIN TYPE-III 2.
FT DOMAIN 494 579 FIBRONECTIN TYPE-III 3.
FT DOMAIN 597 680 FIBRONECTIN TYPE-III 4.
FT DOMAIN 910 1141 PROTEIN-TYROSINE PHOSPHATASE 1.
FT DOMAIN 1200 1433 PROTEIN-TYROSINE PHOSPHATASE 2.
FT ACT_SITE 1082 1082 BY SIMILARITY.
FT ACT_SITE 1376 1376 BY SIMILARITY.
FT DISULFID 216 270 POTENTIAL.
FT SITE 641 644 CLEAVAGE SITE (PROBABLE).
FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 211 211 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 416 416 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 424 424 N-LINKED (GLCNAC. . .) (POTENTIAL).
```



```
FT CARBOHYD 436 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 462 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 552 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 586 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 590 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 690 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 9 L -> V (IN REF. 2).
FT CONFLICT 158 S -> T (IN REF. 2).
FT CONFLICT 284 A -> P (IN REF. 2).
FT CONFLICT 422 T -> S (IN REF. 2).
FT CONFLICT 672 AEL -> CRT (IN REF. 2).
FT CONFLICT 715 S -> T (IN REF. 2).
FT CONFLICT 732 A -> AA (IN REF. 2).
FT CONFLICT 1366 E -> K (IN REF. 2).
SQ SEQUENCE 1439 AA; 162087 MW; EED529AF7C9F4451 CRC64;

Query Match 13.6%; Score 492.5; DB 1; Length 1439;
Best Local Similarity 29.1%; Pred. NO. 1.3e-22;
Matches 136; Conservative 77; Mismatches 167; Indels 87; Gaps 16;

QY 21 KEEFANEFLKRSQTKYKADYTTVAENAKNTKKRYKDLIDYDSRVLSLITSD 80
DB 888 KEYESFF---EGQSASW-----DVAKKQDNRAKNRYGNIAYDHSRVILQPVDDP 936
QY 81 DSSYINANFIKGVGPKAYIATQGPLSTLLDFWRMIWEYSVLIIWACMEYEMGKKCE 140
DB 937 SSDYINANYIDGYQRPVSHYIATQGPVHEVYDFWRMIWQESACIVMTNLVEGVRKCY 996
QY 141 RYWAEPGEMQLFEGPFSVCEAEKRSYIIR--TLKFKNSERTIYQHYKNPDRHDV 198
DB 997 KYWPDDEV---YGFVKVTCVEMELAEYVVRTFLRRGYNEIREVKQFHTGMPDHGV 1053
QY 199 PSSIDPIELIWDVRCYQEDDSVPICHSACGRTGVCALVDYTWMLLKDGIIPEFNS 258
DB 1054 PYHATGLLSFIRVKLSNPPSPAGVIVHCSAGAGRTGCVYIVIDIMDMAERGVV---D 1109
QY 259 VESLREMRTPSLVQTOEQVELYVNAVLEL-----FK-RQMDVIR----- 299
DB 1110 IYNCVKALRSRINNVQVEEQYIFTHDAILEACLGETAIPVCEPKAAYFDMIRDSQTN 1169
QY 300 DKHSTESQAKH-----CIPEKNHTIQ--ADSYSPN--LPKSTTKAAKMMN 341
DB 1170 SSHLKDEFOTLNSVTPRLQAEDCSIACLP-RNHDKNRMDMLPPDRCLPLFLITIDGSSN 1228
QY 342 QORTKM--EIKESSFDFTSISAKEELVHPAKSSTSFDFLELYSF----- 388
DB 1229 YINAALMDSYRQPAAF-----IVTQYPLPNTVKDFWRLVYDYGCTSIYMLNEV 1276
QY 389 DKNADTTMKWQTKAF-----PIVGEPLQKHSLDGLSLFEGCSNSKP 431
DB 1277 DLSQCPQWPEGMRLRGPIQVECMSCSMDCDVINRIFRINLRP 1323

RESULT 10
ID PTN6 HUMAN
AC P29350; STANDARD; PRT; 595 AA.
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Protein-tyrosine phosphatase, non-receptor type 6 (EC 3.1.3.48)
DE (Protein-tyrosine phosphatase 1C) (PTP-1C) (Hematopoietic cell
DE protein-tyrosine phosphatase) (SH-PTP1) (Protein-tyrosine phosphatase
DE SHP-1).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (LONG ISOFORM).
```

```
RX MEDLINE=92123209; PubMed=1732748;
RA Yi T., Cleveland J.L., Ihle J.N.;
RT "Protein tyrosine phosphatase containing SH2 domains:
RT characterization, preferential expression in hematopoietic cells, and
RT localization to human chromosome 12p12-pl3.";
RL Mol. Cell. Biol. 12:836-846(1992).
RN [2]
RP SEQUENCE FROM N.A. (SHORT ISOFORM).
RC TISSUE=Breast;
RX MEDLINE=91343005; PubMed=1652101;
RA Shen S.H., Bastien L., Posner B.I., Chretien P.;
RT "A protein-tyrosine phosphatase with sequence similarity to the SH2
RT domain of the protein-tyrosine kinases.";
RL Nature 352:736-739(1991).
RN [3]
RP REVISIONS.
RA Shen S.H., Bastien L., Posner B.I., Chretien P.;
RL Nature 353:868-868(1991).
RN [4]
RP SEQUENCE FROM N.A. (LONG ISOFORM).
RX MEDLINE=92141214; PubMed=1736296;
RA Plutzky J., Neel B.G., Rosenberg R.D.;
RT "Isolation of a src homology 2-containing tyrosine phosphatase.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:1123-1127(1992).
RN [5]
RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
RX MEDLINE=95394454; PubMed=7665165;
RA Banville D., Stocco R., Shen S.H.;
RT "Human protein tyrosine phosphatase 1C (PTPN6) gene structure:
RT alternate promoter usage and exon skipping generate multiple
RT transcripts.";
RL Genomics 27:165-173(1995).
RN [6]
RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
RX MEDLINE=96303695; PubMed=8723724;
RA Ansari-Lari M.A., Muzny D.M., Lu J., Lu F., Lilley C.E.,
RA Spanos S., Malley T., Gibbs R.A.;
RT "A gene-rich cluster between the CD4 and triosephosphate isomerase
RT genes at human chromosome 12p13.";
RL Genome Res. 6:314-326(1996).
RN [7]
RP PHOSPHORYLATION.
RX MEDLINE=95300784; PubMed=7781604;
RA Li R.Y., Gails F., Ragab A., Ragab-Thomas J.M.F., Chap H.;
RT "Tyrosine phosphorylation of an SH2-containing protein tyrosine
RT phosphatase is coupled to platelet thrombin receptor via a pertussis
RT toxin-sensitive heterotrimeric G-protein.";
RL EMBO J. 14:2519-2526(1995).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 248-399.
RX MEDLINE=98447672; PubMed=9774441;
RA Yang J., Liang X., Niu T., Meng W., Zhao Z., Zhou G.W.;
RT "Crystal structure of the catalytic domain of protein-tyrosine
RT phosphatase SHP-1.";
RL J. Biol. Chem. 273:28199-28207(1998).
CC -!- FUNCTION: PTYS A KEY ROLE IN HEMATOPOIESIS. THIS PTPASE ACTIVITY
CC MAY DIRECTLY LINK GROWTH FACTOR RECEPTORS AND OTHER SIGNALING
CC PROTEINS THROUGH PROTEIN-TYROSINE PHOSPHORYLATION. THE SH2 REGIONS
CC MAY INTERACT WITH OTHER CELLULAR COMPONENTS TO MODULATE ITS OWN
CC PHOSPHATASE ACTIVITY AGAINST INTERACTING SUBSTRATES.
CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O ~ protein
CC tyrosine + phosphate.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE)
CC AND A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN HEMATOPOIETIC
CC CELLS.
CC -!- PTM: PHOSPHORYLATED ON SERINE AND TYROSINE.
CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. SUBCLASS THAT CONTAINS SH2 DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
-----
```


DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR000242; Tyr_prot_phptase.
DR Pfam: PF00194; carb_anhydase; 1.
DR Pfam: PF00041; fn3; 1.
DR Pfam: PF00102; Y_phosphatase; 2.
DR PRINTS: PR00700; PRTYPHPTASE.
DR ProDom: PD000865; Carb_anhydase; 1.
DR SMART: SM00194; PTPC; 2.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS00056; TYR_PHOSPHATASE_2; 2.
DR PROSITE: PS00055; TYR_PHOSPHATASE_PTP; 2.
KW Glycoprotein; Transmembrane; Hydrolase; Repeat; Signal.
FT SIGNAL 1 19
FT CHAIN 20 1442
FT DOMAIN 20 733
FT TRANSMEM 734 759
FT DOMAIN 760 1442
FT DOMAIN 56 322
FT DOMAIN 347 441
FT DOMAIN 866 1122
FT DOMAIN 1123 1442
FT ACT_SITE 1057 1057
FT SITE 1348 1348
FT CARBOHYD 109 109
FT CARBOHYD 113 113
FT CARBOHYD 156 156
FT CARBOHYD 359 359
FT CARBOHYD 444 444
FT CARBOHYD 719 719
SQ SEQUENCE 1442 AA; 161242 MW; 5887715568BECDB CRC64;

Query Match 13.4%; Score 485; DB 1; Length 1442;
Best Local Similarity 25.88; Pred. No. 3.8e-22;
Matches 160; Conservative 96; Mismatches 263; Indels 100; Gaps 19;

Qy 9 KFLDEAQQSKITKEEFANFLKLRQSTKYKADKTYPTTVAENAKNIKKNRYKDILPYD 68
Db 832 KHIGELYSN--SQGFSEDFEEVQRC-----ADNNTAHSNHPDNKKNRYINLAYDH 885

Qy 69 SRVSLSLTSD--SSYINANFTKGVGPKAYATOGPLSTLLDFWRMIWEYSVLIV 126
Db 886 SRVLRPLPGKDSRHSYINANYVDGYNKAKYATOGPLKSTPDEFWRMIWEQNTGIII 945

Qy 127 MACMEYEMGKKCRYYAEPCEMOLRFPFSVCEAEKRKSDYIIRTKV----- 176
Db 946 MITNLVEGRKRKCDQY--PTENTEENIIVITUKSTKVHACYTVRRLSVNRTKVKKGQK 1003

Qy 177 ---RFSNSETRTIYOFHYKNWPDHDPSSIDPILSLIWDVRCYQEDDSVPICIHCSAGCGR 233
Db 1004 GNPGRQNERTVIOYHTQWDMGVEYALPVLTFVRRSSAARMPPDGMPLVHCSAGVR 1063

Qy 234 TGVTCIVDYTWMLLKQGITPENSFVSLIREMTQPSLVQTOEQYELVYNVAVLEFKR 293
Db 1064 TGTY-IVTDSMLQIKD---KSTVNLGFLKHIRTQRYNLVQTEEQYIFIHIDALLE---- 1115

Qy 294 QMDVIRKHSCTESQAKH-----CIPEKNTLQADSVSPNLKSTTKAAKMMNQORTKM 347
Db 1116 ---ALIGKETAVSSQLSHSYNSILIPGVGCKTRLEQFKLITOCNAKYVBCFSAQ---- 1168

Qy 348 EIKESSSDFRTSEI--SAKEELVHLPAKSTSFDFLELANSFQKNADTTMKWTKAPFI 405
Db 1169 --KECNKERNSSVPAERARVGLAPLPGMKGTDIYNASYIMG-----YYSRNEFII 1219

Qy 406 VGEPLQKHQSLDLSGLFEGGCSNKPVNAAGRYFNSKVPITRTKSTPELLQOQRETKVD 465
Db 1220 TQHPD-LPHTTKDFRMWDH-----NA-----QIIVMLPDNQSL 1252

Qy 466 SKENFYSLEPHDSCEFEMOAKVMHVSSAELNYSLPDYSKHOIRNASNVKHHDSALG 525
Db 1253 AEDEFVYWPSSREES---MNCFAFTVTLISKDRCLLS-----NEEQIIIHDFILEA 1299

Qy 526 VY-SYIPLVENPYFSSWPPSGTSSKMSLDLPEKQDGTVPFSSLLPTSSLSFYNSHSS 584
Db 1300 TODDYVLEVRHFQCPKPNPDAPISSTFEL-----INVKEALTRDGPPTIVHDEYGA 1355

Qy 585 LSLNSPTNISLLNQESAV 603
Db 1356 GMLCALTTLSQLENAV 1374

RESULT 12
PTPM_MOUSE STANDARD; PRT; 1452 AA.
AC P28828;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein-tyrosine phosphatase MU precursor (EC 3.1.3.48) (R-PTP-MU).
GN PTPRM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=92008644; PubMed=1655529;
RA Gebbink M.F.B.G., van Etten I., Hateboer G., Suijkerbuijk R.,
RA Beijersbergen R., Geurts van Kessel A., Moolenaar W.H.;
RT "Cloning, expression and chromosomal localization of a new putative
RT receptor-like protein tyrosine phosphatase";
RL FEBS Lett. 290:123-130(1991).
CC -!- FUNCTION: MAY PLAY A KEY ROLE IN SIGNAL TRANSDUCTION AND GROWTH
CC CONTROL.
CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: MOST ABUNDANT IN LUNG, LESS IN BRAIN AND
CC HEART.
CC -!- SIMILARITY: CONTAINS 1 MAM DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC -!- SIMILARITY: CONTAINS 4 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X58287; CAA41225.1; -.
CC PIR: S17670; S17670.
CC HSP: P28827; IRPM.
CC MGD: MGI:102694; Pcprrm.
CC InterPro: IPR003961; FN_III.
CC InterPro: IPR003962; FNIII_repeat.
CC InterPro: IPR003599; Ig.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR000998; MAM.
CC InterPro: IPR000387; Tyr_phosphatase.
CC InterPro: IPR000242; Tyr_prot_phptase.
CC Pfam: PF00041; fn3; 3.
CC Pfam: PF00047; Ig; 1.
CC Pfam: PF00629; MAM; 1.
CC Pfam: PF00102; Y_phosphatase; 2.
CC PRINTS: PR00014; FNTYPEIII.
CC PRINTS: PR00020; MAMDOMAIN.
CC PRINTS: PR00700; PRTYPHPTASE.
CC SMART: SM00060; FN3; 2.
CC SMART: SM00409; IG; 1.
CC SMART: SM00137; MAM; 1.
CC SMART: SM00194; PTPC; 2.

[3]
RC SEQUENCE OF 1479-2091 FROM N.A. (LONG FORM).
RC TISSUE=Liver;
RX MEDLINE=91006018; PubMed=2170109;
RA Krueger N.X., Streuli M., Saito H.;
RT "Structural diversity and evolution of human receptor-like protein
tyrosine phosphatases";
RL EMO J. 9:3241-3252(1990).
[4]
RP SEQUENCE OF 1749-1990 AND 2047-2280 FROM N.A. (LONG FORM).
RC TISSUE=Brain stem; PubMed=2169617;
RX MEDLINE=90384936; PubMed=2169617;
RA Kaplan R., Morse B., Huebner K., Croce C., Howk R., Ravera M.,
RA Ricca G., Jaye M., Schlessinger J.;
RT "Cloning of three human tyrosine phosphatases reveals a multigene
family of receptor-linked protein-tyrosine-phosphatases expressed in
brain";
RT proc.;
RL Abn. Natl. Acad. Sci. U.S.A. 87:7000-7004(1990).
CC -!- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF SPECIFIC
DEVELOPMENTAL PROCESSES IN THE CNS.
CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
tyrosine + phosphate.
CC -!- SUBUNIT: THE CARBONIC-ANHYDRASE LIKE DOMAIN BINDS TO CONTACTIN (BY
SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN THE CENTRAL NERVOUS
SYSTEM, WHERE IT IS LOCALIZED IN THE PURKINJE CELL LAYER OF THE
CEREBELLUM, THE DENTATE GYRUS, AND THE SUBEPENDYMAL LAYER OF THE
ANTERIOR HORN OF THE LATERAL VENTRICLE. DEVELOPMENTALLY REGULATED
IN THE BRAIN.
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE EUKARYOTIC-
TYPE CARBONIC ANHYDRASE FAMILY.
CC -!- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
CC -!- CAUTION: CALLED RPPASE BETA IN REF.2 AND 4.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; M93426; AAA60225.1; -;
DR EMBL; X54135; CAA38070.1; -;
DR PIR; S12054; S12054.
DR PIR; A46151; A46151.
DR HSSP; P18052; 1YFO.
DR MIM; 176891; -;
DR InterPro; IPR001148; Carb.anhydrase.
DR InterPro; IPR003961; FN.III.
DR InterPro; IPR000387; Tyr.phosphatase.
DR InterPro; IPR000242; Tyr_prot_phptase.
DR Pfam; PF00194; carb.anhydrase; 1.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00102; Y_phosphatase; 2.
DR PRINTS; PR00700; PTRYPHPTASE.
DR PRODOM; PD000865; Carb.anhydrase; 1.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00194; PTPC; 2.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 2.
DR PROSITE; PS00055; TYR_PHOSPHATASE_PTP; 2.
KW Signal; Glycoprotein; Transmembrane; Hydrolase; Repeat;
KW Alternative splicing.
FT SIGNAL 1 24
FT CHAIN 25 2314
FT DOMAIN 25 1635
FT TRANSMEM 1636 1661
FT DOMAIN 1662 2314
BY SIMILARITY.
FT FT
FT DOMAIN 25 2314
FT TRANSMEM 1636 1661
FT DOMAIN 1662 2314
CYTOPLASMIC (POTENTIAL).

DOMAIN 34 302
FT DOMAIN 312 406
FT DOMAIN 1744 1997
FT DOMAIN 1998 2314
FT BINDING 587 587
FT BINDING 637 637
FT BINDING 997 997
FT BINDING 1548 1548
FT BINDING 1550 1550
FT ACT_SITE 1932 1932
FT SITE 2222 2222
FT CARBOHYD 105 105
FT CARBOHYD 134 134
FT CARBOHYD 223 223
FT CARBOHYD 232 232
FT CARBOHYD 324 324
FT CARBOHYD 381 381
FT CARBOHYD 497 497
FT CARBOHYD 501 501
FT CARBOHYD 552 552
FT CARBOHYD 602 602
FT CARBOHYD 629 629
FT CARBOHYD 677 677
FT CARBOHYD 1017 1017
FT CARBOHYD 1050 1050
FT CARBOHYD 1082 1082
FT CARBOHYD 1122 1122
FT CARBOHYD 1456 1456
FT CARBOHYD 1561 1561
FT CARBOHYD 1617 1617
FT VARSPIC 755 1614
FT CONFLICT 1722 1728
SQ SEQUENCE 2314 AA; 254528 MW; 77DBDEF4A0F5FB42 CRC64;
CARBONIC-ANHYDRASE LIKE.
FIBRONECTIN TYPE-III.
PROTEIN-TYROSINE PHOSPHATASE 1.
PROTEIN-TYROSINE PHOSPHATASE 2.
CHONDROITIN SULFATE (POTENTIAL).
CHONDROITIN SULFATE (POTENTIAL).
CHONDROITIN SULFATE (POTENTIAL).
CHONDROITIN SULFATE (POTENTIAL).
BY SIMILARITY.
ANCESTRAL ACTIVE SITE.
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
MISSING (IN SHORT ISOFORM).
MISSING (IN REF. 2).
Query Match 13.2%; Score 478; DB 1; Length 2314;
Best Local Similarity 26.6%; Pred. No. 2e-21;
Matches 140; Conservative 83; Mismatches 192; Indels 112; Gaps 17;
Qy 2 DORETLOQLFDEAOSKKTKEEFANEFLLKROSKYKADKTYPTVAENAKNIKKNRYK 61
Db 1718 EEFELKFEYQEVQSCYVD-----LGITADSSNHDPNK-----HKNYI 1756
Qy 62 DILPYDYSRVELSLTSDDED--SSYINANFTKGVYGPXKAYIATQGPLSTLLDFWRMIWE 119
Db 1757 NIVAYDHSRVKLAQLAEXDKGLTDYINANYVDYGNRPXAYIAAQGLKSTAEDEFWRMIWE 1816
Qy 120 YSVLIIVNACMEYEMGKKKCYRYHAEPGEMOLEFCPP-----SVSCEAEKRKSDYIIRTL 174
Db 1817 HNVEVIVMTNLVKEGRRKCDQYW--PADGSEYGNFLVTKQSVQVLAYYTVRNFTLRNT 1874
Qy 175 KVKENSET-----RTIYQFYHKNWPDHVPSSIDPILLEIWDVRCYQEDDSVPICIHCSA 229
Db 1875 KIKGSGQRSPGRVVTQYHYTQWDMGVPEYSLPVLFFVRKAAKRAHVGVPVVHCSA 1934
Qy 230 CGGRGTGVCIAVDYTWMLLKDGIIIPENSVFLIREMTQRPSLVQTQEQVELVYNAVLE 289
Db 1935 GVGRTGTIVLDSMLQIQIHGTV---NIFGLKHRSQRNLYLVQTEQYVFIHDTLVE 1990
Qy 290 LFKROMDVIRKHX-----SGTESQAKICIEKKNHTLOADSYSP-----NLPK 331
Db 1991 AILSKETEVLDSHIHAYVYNALLIPGPAKGTLEKQFOLLSSQNSIQOOSYSAALKOCNREK 2050
Qy 332 STTKAAKMMNQORTKMEIKESSSF-----FRTSEISAKELVHLPKASSTSF 381
Db 2051 NRTSIIIPVRSRGVIGISLSGEGTDYINASYMGVYQNEFIITQHPLLHTIK-----DF 2105
Qy 382 LELNYSFDKNA-----DTTMKWOTKAFPIVGE-----LOKHQS 415
Db 2106 WRM--IWDHNAQLVVMIPDQGNMAEDFVYFNKDPNCESFKVTLMAEEHKCLSNREK 2163
Qy 416 LDGLSLLFEGGSNSKPVNAAGRYF-----NSKVPITRTKSTPFFELI 456
Db 416 LDGLSLLFEGGSNSKPVNAAGRYF-----NSKVPITRTKSTPFFELI 456

Db 2164 LIIQDFILEATQDDYVLEV--RHFQCPKWPNDSPISKT-----FELI 2204

Search completed: May 26, 2002, 12:26:53
Job time: 538 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2002, 12:16:55 ; Search time 95.55 Seconds
(without alignments)
1252.878 Million cell updates/sec

Title: US-09-600-358A-4
Perfect score: 3615
Sequence: 1 MDQREILQKFLDEAQSCKIT.....RPSKSVKLRSRKGKNSWL 692

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3611	99.9	692	4	O95064
2	3569	98.7	808	4	O95063
3	3512.5	97.2	807	4	O9Y2R2
4	2971.5	82.2	799	4	O9P0U2
5	1042	28.8	382	11	O63745
6	748	20.7	453	11	P70602
7	735.5	20.3	453	11	O922E3
8	733.5	20.3	453	11	O61152
9	708	19.6	458	4	O99952
10	515.5	14.3	1114	4	O9H0V6
11	515.5	14.3	1143	4	Q16614
12	513	14.2	361	11	O61373
13	511	14.1	1195	13	O9PUM2
14	508	14.1	1216	11	O62884
15	508	14.1	1337	4	Q15255
16	507	14.0	1291	11	Q61812

17	498	13.8	492	4	O75664
18	498	13.8	1435	11	O9JIZ2
19	498	13.8	1445	11	O99M81
20	498	13.8	1454	11	O99M82
21	498	13.8	1463	4	O43655
22	494	13.7	1444	11	O9JIZ1
23	492	13.6	1444	11	O9JLP0
24	490	13.6	1216	13	O9IBD8
25	488.5	13.5	2271	13	O9I909
26	487.5	13.5	832	13	O90947
27	485	13.4	597	4	O969V8
28	483.5	13.4	593	13	O90687
29	483.5	13.4	1200	13	Q91054
30	482	13.3	624	4	O9UK67
31	480.5	13.3	487	5	O9NL13
32	479	13.3	1455	11	Q99M80
33	478	13.2	1648	4	O9UDR6
34	477.5	13.2	1788	13	O9IAJ0
35	477	13.2	613	11	P81718
36	475.5	13.2	508	13	O90VJ5
37	475.5	13.2	1237	13	O91976
38	474.5	13.1	468	13	O9IBAS
39	474.5	13.1	1115	4	O9HD43
40	473	13.1	470	5	O9NL15
41	473	13.1	1406	13	Q9W6V5
42	472.5	13.1	438	11	O64642
43	471.5	13.0	597	11	O64509
44	471	13.0	1214	5	O9VAL3
45	470.5	13.0	468	13	Q9IBAS

ALIGNMENTS

RESULT 1

O95064 PRELIMINARY; PRT; 692 AA.
 AC O95064;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE LYPHOID PHOSPHATASE LYP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_taxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Roifman C.M.;
 RT "cDNA of Lyp2 Protein Tyrosine Phosphatase."
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF001847; AAD00905.1;
 DR HSSP; Q06124; 2SHP.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR InterPro; IPR000242; TYR_prot_phptase.
 DR Pfam; PF00102; Y_phosphatase; 1.
 DR PRINTS; PR00700; PRTYPHPTASE.
 DR SMART; SM00194; PTPc; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS00055; TYR_PHOSPHATASE_ptp; 1.
 KW Hydrolase.
 SQ SEQUENCE 692 AA; 78828 MW; 543ED06AE9A8C4DF CRC64;

Query Match 99.9%; Score 3611; DB 4; Length 692;

Best Local Similarity 99.9%; Pred. No. 5.9e-249;

Matches 691; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MDQREILQKFLDEAQSCKITKEEFANEFKLKRQSTYKADKYPTTVAENAKIKNNRY 60

Db 1 MDQREILQKFLDEAQSCKITKEEFANEFKLKRQSTYKADKYPTTVAENAKIKNNRY 60

QY 61 KDILPYDSRVLSLITSDSSSYINANFIKGVYGPAYIATQGPLSTLLDFWRMIWEY 120
Db 61 KDILPYDSRVLSLITSDSSSYINANFIKGVYGPAYIATQGPLSTLLDFWRMIWEY 120
QY 121 SVLIIVMACMEYEMGKKKERYWAEPPGEMQLEFGFPFVSCAEKRKSDYIIRTLKVKENS 180
Db 121 SVLIIVMACMEYEMGKKKERYWAEPPGEMQLEFGFPFVSCAEKRKSDYIIRTLKVKENS 180
QY 181 ETRTIYQHYKNWPDHVPSSIDPILWVRCYQEDDSVPICIHCSAGCGRTGVCICAI 240
Db 181 ETRTIYQHYKNWPDHVPSSIDPILWVRCYQEDDSVPICIHCSAGCGRTGVCICAI 240
QY 241 VDYTWMLLKDGIIIPENFVSFLIREMRTQPSLVOTQEQYELVYNAVLELFRQMDVIRD 300
Db 241 VDYTWMLLKDGIIIPENFVSFLIREMRTQPSLVOTQEQYELVYNAVLELFRQMDVIRD 300
QY 301 KHSGETESQAKHCIPKKNHTLQADSYSPNLPKSTTTAAKMMNQORTKMEIKESSSDFDRTS 360
Db 301 KHSGETESQAKHCIPKKNHTLQADSYSPNLPKSTTTAAKMMNQORTKMEIKESSSDFDRTS 360
QY 361 EISAKEELVLHPAKSSTSFDFLELNYSPDKNADTTMKWTKAFPIVGPLOKHQSILDLGS 420
Db 361 EISAKEELVLHPAKSSTSFDFLELNYSPDKNADTTMKWTKAFPIVGPLOKHQSILDLGS 420
QY 421 LFFEGCSNSKPVNAAGRYFNSKVPITRTKSTPFELIQORETKEVDKSNFYLESQPHDS 480
Db 421 LFFEGCSNSKPVNAAGRYFNSKVPITRTKSTPFELIQORETKEVDKSNFYLESQPHDS 480
QY 481 CFVEMQAKQVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGVYIPLVNPYFSS 540
Db 481 CFVEMQAKQVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGVYIPLVNPYFSS 540
QY 541 WPPSGTSSKMSLDLPEKODGTVPFSSLLPTSTSLFSYNSHSSLSLNSPTNISLLNQE 600
Db 541 WPPSGTSSKMSLDLPEKODGTVPFSSLLPTSTSLFSYNSHSSLSLNSPTNISLLNQE 600
QY 601 SAVLATAPRIDDEIPPLPVRTPEFIVVEEAGEFSPNVPKSLSSAVKVKIGTSLWGGT 660
Db 601 SAVLATAPRIDDEIPPLPVRTPEFIVVEEAGEFSPNVPKSLSSAVKVKIGTSLWGGT 660
QY 661 SEPCKFDDSVILRPSKSVKLRSPKS 685
Db 661 SEPCKFDDSVILRPSKSVKLRSPKS 685

RESULT 2
Q95063 PRELIMINARY; PRT; 808 AA.
ID O95063
AC O95063;
DT 01-MAY-1999 (TremBLrel. 10, Created)
DT 01-MAY-1999 (TremBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE LYMPHOID PHOSPHATASE LypI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Roifman C.M.;
RT "Human cDNA of LypI protein Tyrosine Phosphatase.";
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF001846; AAD00904.1; -
DR HSP: Q06124; 2SHP.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR000242; Tyr_prot_phptase.
DR Pfam: PF00102; Y_phosphatase; 1.
DR PRINTS: PR00700; PRTYPHPTASE.
DR SMART: SM00194; PTPC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase.

SQ SEQUENCE 808 AA; 91763 MW; 8375848DF8D3D853 CRC64;
Query Match 98.7%; Score 3569; DB 4; Length 808;
Best Local Similarity 99.9%; Pred. No. 7.le-246;
Matches 684; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDQREILQKFDEAQSKKITKEEFANEFKLKRSQTKYKADKTYPTTVAENAKNKKRY 60
Db 1 MDQREILQKFDEAQSKKITKEEFANEFKLKRSQTKYKADKTYPTTVAENAKNKKRY 60
QY 61 KDILPYDSRVLSLITSDSSSYINANFIKGVYGPAYIATQGPLSTLLDFWRMIWEY 120
Db 61 KDILPYDSRVLSLITSDSSSYINANFIKGVYGPAYIATQGPLSTLLDFWRMIWEY 120
QY 121 SVLIIVMACMEYEMGKKKERYWAEPPGEMQLEFGFPFVSCAEKRKSDYIIRTLKVKENS 180
Db 121 SVLIIVMACMEYEMGKKKERYWAEPPGEMQLEFGFPFVSCAEKRKSDYIIRTLKVKENS 180
QY 181 ETRTIYQHYKNWPDHVPSSIDPILWVRCYQEDDSVPICIHCSAGCGRTGVCICAI 240
Db 181 ETRTIYQHYKNWPDHVPSSIDPILWVRCYQEDDSVPICIHCSAGCGRTGVCICAI 240
QY 241 VDYTWMLLKDGIIIPENFVSFLIREMRTQPSLVOTQEQYELVYNAVLELFRQMDVIRD 300
Db 241 VDYTWMLLKDGIIIPENFVSFLIREMRTQPSLVOTQEQYELVYNAVLELFRQMDVIRD 300
QY 301 KHSGETESQAKHCIPKKNHTLQADSYSPNLPKSTTTAAKMMNQORTKMEIKESSSDFDRTS 360
Db 301 KHSGETESQAKHCIPKKNHTLQADSYSPNLPKSTTTAAKMMNQORTKMEIKESSSDFDRTS 360
QY 361 EISAKEELVLHPAKSSTSFDFLELNYSPDKNADTTMKWTKAFPIVGPLOKHQSILDLGS 420
Db 361 EISAKEELVLHPAKSSTSFDFLELNYSPDKNADTTMKWTKAFPIVGPLOKHQSILDLGS 420
QY 421 LFFEGCSNSKPVNAAGRYFNSKVPITRTKSTPFELIQORETKEVDKSNFYLESQPHDS 480
Db 421 LFFEGCSNSKPVNAAGRYFNSKVPITRTKSTPFELIQORETKEVDKSNFYLESQPHDS 480
QY 481 CFVEMQAKQVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGVYIPLVNPYFSS 540
Db 481 CFVEMQAKQVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGVYIPLVNPYFSS 540
QY 541 WPPSGTSSKMSLDLPEKODGTVPFSSLLPTSTSLFSYNSHSSLSLNSPTNISLLNQE 600
Db 541 WPPSGTSSKMSLDLPEKODGTVPFSSLLPTSTSLFSYNSHSSLSLNSPTNISLLNQE 600
QY 601 SAVLATAPRIDDEIPPLPVRTPEFIVVEEAGEFSPNVPKSLSSAVKVKIGTSLWGGT 660
Db 601 SAVLATAPRIDDEIPPLPVRTPEFIVVEEAGEFSPNVPKSLSSAVKVKIGTSLWGGT 660
QY 661 SEPCKFDDSVILRPSKSVKLRSPKS 685
Db 661 SEPCKFDDSVILRPSKSVKLRSPKS 685
RESULT 3
Q9Y2R2 PRELIMINARY; PRT; 807 AA.
ID Q9Y2R2
AC Q9Y2R2;
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE PROTEIN TYROSINE PHOSPHATASE HOMOLOG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu T., Zhang J., Fu G., Zhang Q., Ye M., Zhou J., Wu J., Shen Y.,
RA Yu M., Chen S., Mao M., Chen Z.;
RT "Human protein tyrosine phosphatase (70zpep) homolog.";

RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL: A0707031; AAD2764.1; -;
DR HSSP: Q06124; 2SHP.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR000242; Tyr_prot_phptase.
DR Pfam: PF00102; Y_phosphatase; 1.
DR PRINTS: PR00700; PRTVPHPTASE.
DR SMART: SM00194; PTPC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase.
SQ SEQUENCE 807 AA; 91708 MW; A5B4B4CAB795B812 CRC64;

Query Match 97.2%; Score 3512.5; DB 4; Length 807;
Best Local Similarity 98.8%; Pred. No. 7.5e-242;
Matches 677; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 1 MDQREILQKFLDEAQSCKITKEEFANFLKQSTKYKADKTYPTTVAENAKNKKRY 60
DB 1 MDQREILQKFLDEAQSCKITKEEFANFLKQSTKYKADKTYPTTVAEKPNKKRY 60
QY 61 KDILPYDSRVLSLTSDESSYINANFIKGYGPKAYIATQGPLSTTLDFWRMIWEY 120
DB 61 KDILPYDSRVLSLTSDESSYINANFIKGYGPKAYIATQGPLSTTLDFWRMIWEY 120
QY 121 SVLIIVMACMEYEMGKKKERYWAEFGEMQLEFGFVSCEAEKRSKSDYIIRTLVKVFN 180
DB 121 SVLIIVMACMEYEMGKKKERYWAEFGEMQLEFGFVSCEAEKRSKSDYIIRTLVKVFN 180
QY 181 ERTTIYQFHYKKNPDHVPSSIDPILLELWDVRCQEDDSVPICHCAGCGRTGVICAI 240
DB 181 ERTTIYQFHYKKNPDHVPSSIDPILLELWDVRCQEDDSVPICHCAGCGRTGVICAI 240
QY 241 VDTYTMLLKDGIIIPENFVSFLIREMRTQPSLVQTQEQYELVYNAVLELFRQMDVIRD 300
DB 241 -DVTYTMLLKDGIIIPENFVSFLIREMRTQPSLVQTQEQYELVYNAVLELFRQMDVIRD 299
QY 301 KHSGETSQAKHCIPKKNHTLQADSYSPNLPKSTTTAAKMMNQORTKMEIKESSDFRTS 360
DB 300 KHSGETSQAKHCIPKKNHTLQADSYSPNLPKSTTTAAKMMNQORTKMEIKESSDFRTS 359
QY 361 ELSAKEELVHPAKSTSFDFLELNYSPDKNADTTMKWTKAFPIVGEPLQKHQSLDLS 420
DB 360 ELSAKEELVHPAKSTSFDFLELNYSPDKNADTTMKWTKAFPIVGEPLQKHQSLDLS 419
QY 421 LLEFGCSNKPVNAAGRVNSKVPITRTKSTPFELIQORETREVDSKENSFYLESO 480
DB 420 LLEFGCSNKPVNAAGRVNSKVPITRTKSTPFELIQORETREVDSKENSFYLESO 479
QY 481 CFVEMQAQKMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGVYSYIPLVNPYFSS 540
DB 480 CFVEMQAQKMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGVYSYIPLVNPYFSS 539
QY 541 WPPSGTSSKMSLDLPEKQDGVFPSSLLPTSTSLFSYNSHSLSLNSPTNISLLNOE 600
DB 540 WPPSGTSSKMSLDLPEKQDGVFPSSLLPTSTSLFSYNSHSLSLNSPTNISLLNOE 599
QY 601 SAVLATAPRIDDEIIPPLPVRTPEFIVVEEAGEFSPNPKSLSSAVKVKICTSLEWGT 660
DB 600 SAVLATAPRIDDEIIPPLPVRTPEFIVVEEAGEFSPNPKSLSSAVKVKICTSLEWGT 659
QY 661 SEPKFDDSVILRPSKSVLRSPKS 685
DB 660 SEPKFDDSVILRPSKSVLRSPKS 684

RESULT 4

ID Q9P002 PRELIMINARY; PRT; 799 AA.
AC Q9P002;
DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PROTEIN TYROSINE PHOSPHATASE.
GN PTP.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ADRENAL GLAND;
RA Huang C., Zhang C., Wu T., Peng Y., Gu Y., Zhang L., Jiang C., Li Y.,
RA Han Z., Wang Y., Chen Z., Fu G.;
RT "A novel gene expressed in human adrenal gland";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR HSSP: Q06124; 2SHP.
DR EMBL: AF150732; AAF67472.1; -;
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR000242; Tyr_prot_phptase.
DR PRINTS: PR00700; PRTVPHPTASE.
DR SMART: SM00194; PTPC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase.
SQ SEQUENCE 799 AA; 90610 MW; 93F5385016F33D0C CRC64;

Query Match 82.2%; Score 2971.5; DB 4; Length 799;
Best Local Similarity 87.3%; Pred. No. 2.6e-203;
Matches 598; Conservative 17; Mismatches 61; Indels 9; Gaps 9;

QY 1 MDQREILQKFLDEAQSCKITKEEFANFLKQSTKYKADKTYPTTVAENAKNKKRY 60
DB 1 MDQREILQKFLDEAQSCKITKEEFANFLKQSTKYKADKTYPTTVAEKPNKKRY 60
QY 61 KDILPYDSRVLSLTSDESSYINANFIKGYGPKAYIATQGPLSTTLDFWRMIWEY 120
DB 61 KDILPYDSRVLSLTSDESSYINANFIKGYGPKAYIATQGPLSTTLDFWRMIWEY 120
QY 121 SVLIIVMACMEYEMGKKKERYWAEFGEMQLEFGFVSCEAEKRSKSDYIIRTLVKVFN 180
DB 121 SVLIIVMACMEYEMGKKKERYWAEFGEMQLEFGFVSCEAEKRSKSDYIIRTLVKVFN 180
QY 181 ERTTIYQFHYKKNPDHVPSSIDPILLELWDVRCQEDDSVPICHCAGCGRTGVICAI 240
DB 181 ERTTIYQFHYKKNPDHVPSSIDPILLELWDVRCQEDDSVPICHCAGCGRTGVICAI 240
QY 241 VDTYTMLLKDGIIIPENFVSFLIREMRTQPSLVQTQEQYELVYNAVLELFRQMDVIRD 300
DB 241 -DVTYTMLLKDGIIIPENFVSFLIREMRTQPSLVQTQEQYELVYNAVLELFRQMDVIRD 299
QY 301 KHSGETSQAKHCIPKKNHTLQADSYSPNLPKSTTTAAKMMNQORTKMEIKESSDFRTS 360
DB 300 KHSGETSQAKHCIPKKNHTLQADSYSPNLPKSTTTAAKMMNQORTKMEIKESSDFRTS 359
QY 361 ELSAKEELVHPAKSTSFDFLELNYSPDKNADTTMKWTKAFPIVGEPLQKHQSLDLS 420
DB 360 ELSAKEELVHPAKSTSFDFLELNYSPDKNADTTMKWTKAFPIVGEPLQKHQSLDLS 416
QY 421 LLEFGCSNKPVNAAGRVNSKVPITRTKSTPFELIQORETREVDSKENSFYLESO 480
DB 421 LLEFGCSNKPVNAAGRVNSKVPITRTKSTPFELIQORETREVDSKENSFYLESO 471
QY 481 CFVEMQAQKMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGVYSYIPLVNPYFSS 540
DB 481 CFVEMQAQKMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGVYSYIPLVNPYFSS 531
QY 541 WPPSGTSSKMSLDLPEKQDGVFPSSLLPTSTSLFSYNSHSLSLNSPTNISLLNOE 600
DB 541 WPPSGTSSKMSLDLPEKQDGVFPSSLLPTSTSLFSYNSHSLSLNSPTNISLLNOE 591
QY 601 SAVLATAPRIDDEIIPPLPVRTPEFIVVEEAGEFSPNPKSLSSAVKVKICTSLEWGT 660

Db 592 SAVLATAPRIDDEIPLHFLXGTSPFVWEAGEFSPNPKSLSSAVKVGKIGTSLWGGT 651
Qy 661 SEPKKFDSDVILRPSKSVKLRSPKS 685
Db 652 SEPKKFDSDVILRPSKSVKLRSPKS 676
RESULT .5
Q63745
ID Q63745 PRELIMINARY; PRT; 382 AA.
AC Q63745;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PROTEIN TYROSINE PHOSPHATASE.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RX MEDLINE=95046282; PubMed=7957881;
RA Moriyama T., Kawanishi S., Inoue T., Inai E., Kaneko T., Xia C.,
RA Takenaka M., Kanada T., Ueda N.;
RT "cDNA cloning of a cytosolic protein tyrosine phosphatase (RKPTP) from
RT rat kidney.";
RL FEBS Lett. 353:305-308(1994).
DR EMBL; D38072; BAA07266.1; -.
DR HSP; Q06124; 2SHP.
DR InterPro; IPR000387; Tyr_phosphatase.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00700; PRTPHPHTASE.
DR SMART; SM00194; PTPc; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase.
SQ SEQUENCE 382 AA; 44438 MW; 25A517FDEBB7C344 CRC64;

Query Match 28.8%; Score 1042; DB 11; Length 382;
Best Local Similarity 53.5%; Pred. No. 2.8e-66;
Matches 201; Conservative 62; Mismatches 89; Indels 24; Gaps 5;
Qy 1 MDQREILQKFLDERAQSKKIT----KEEFANFLKLRQSTKYKADKYPTTVAENAKNIK 56
Db 1 MEQVEILRRFQRYQAMKSPDHNGEDNFARDFMLRLRLSTKYRTEKIYPTATGEKENVK 60
Qy 57 KNRVKDILPYDYSRVLSLITSDSDSSYNANFTKGVYGPAYATATOGPLSTLLDFWRM 116
Db 61 KNRKDLILPFDHRSRVKFLTKTPSQSDYINANFINGVYGPAYATOGPLANTVIDFWRM 120
Qy 117 IWEXSVLIIVMACMEYEMGKKKCYRYAEPGEMQLEFGPPFSVSCAEAKRKS DYIIRLKV 176
Db 121 IWENVVLIIVMACREFEMGRKKCYRYWPLYGEDPITPAFFKISCENEQARTDYFIRTL 180
Qy 177 KFNSETRIYGFHYKNWPDHVPSSIDPILFELIWDVRCYQEDDSVPICIHCSAGCGRTGV 236
Db 181 EFONESRLYGFHYVNWPDHVPSSDFSDILDMISLMRKRYQEHEDVPICIHCSAGCGRTGA 240
Qy 237 ICALIVDYTMLLKDGIIIPENSVFSLIREMTQRPVSLVOTQOYELVYNVNALEFKROMD 296
Db 241 ICAL-DYTMNLLKAGKIPEDFNENLIQEMTQRHSAVQTKEQYELVHRAIAQLFEKQLQ 299
Qy 297 VIRDKHSGTESQAKHCIPKKNHTLOADSYSPNLKPKSTTTKAAKMMNQORTKMEIKESSFD 356
Db 300 -LYEIHGAQKITDGNIEISTGNWSSIDSEKQDSPPP-----KPPRTRSLVEG----- 346
Qy 357 FRTSISAKKEELVLP 372
|||||:|

Db 347 -----DAKEEILQPP 356
RESULT 6
P70602
ID P70602 PRELIMINARY; PRT; 453 AA.
AC P70602;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PROTEIN TYROSINE PHOSPHATASE 20.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97067206; PubMed=8910608;
RA Aoki N., Yamaguchi-Aoki Y., Ullrich A.;
RT "The novel protein-tyrosine phosphatase PTP20 is a positive regulator
RT of PC12 cell neuronal differentiation.";
RL J. Biol. Chem. 271:29422-29426(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Aoki, N, Yamaguchi-Aoki, Y, Ullrich A.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RE EMBL; U69673; AAC52896.1; -.
DR HSP; Q06124; 2SHP.
DR InterPro; IPR000387; Tyr_phosphatase.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00700; PRTPHPHTASE.
DR SMART; SM00194; PTPc; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase.
SQ SEQUENCE 453 AA; 50106 MW; 6CCCI32206FB69AE CRC64;
Query Match 20.7%; Score 748; DB 11; Length 453;
Best Local Similarity 43.1%; Pred. No. 3.1e-45;
Matches 154; Conservative 58; Mismatches 117; Indels 28; Gaps 7;
Qy 3 QREILQKFLDERAQSKKITKEE-FANFELKLRQSTKYKADKYPTTVAENAKNIKKNRYK 61
Db 4 QSDLVRSFLEQGEARDHRKGAILAREFSDIKARSVAMKTEGVCSTKAGSQGSKKNRYK 63
Qy 62 DILPYDYSRVLSLITSDSDSSYNANFTKGVYGPAYATATOGPLSTLLDFWRMWEYS 121
Db 64 DVVPYDETRVTLSSLQEEGHGDIYANFIRCTGDSQAYIATQGPLHTLLDFWRLVWFG 123
Qy 122 VLIIVMACMEYEMGKKKCYRYAEPGEMQLEFGPPFSVSCAEKR-KSDYIIRTLKVKFNS 180
Db 124 IKVILMACQETENGRRKRCERYWAQERE-PLQAGPFCITLTRETALTSIDITURLOVTFQK 182
Qy 181 ETRITRYGFHYKNWPDHVPSSIDPILFELIWDVRCYQEDDSVPICIHCSAGCGRTGVICA 240
Db 183 ESRPVHQLQYVMSWPDHGVPSSSDHLTMVEARCLQGLGPGPLCVHCSAGCGRTGVICA- 241
Qy 241 VDYTMNLLKDGIIIPENSVFSLIREMTQRPVSLVOTQOYELVYNVNALEFKROMD---- 296
Db 242 VDYVROLQLLTQITIPNFSLFVWLEMRKQRPAAVQTREEQYRFLYHTVAQLFSRTLNNSP 301
Qy 297 -----VIRDKHSGTESQAKHCIPKKNHTLOADSYSPNLKPKSTTKA 336
Db 302 LYQNLKENRAPICKDSSSLRTSSA---LPATSRPLGGVLRISVPGPTPLMADTYA 355
RESULT 7
Q922E3
ID Q922E3 PRELIMINARY; PRT; 453 AA.
AC Q922E3;

Db 301 PHYONKENCAPICKEAFSLRTSSALP 327

RESULT 9
Q99952 ID Q99952 PRELIMINARY; PRT; 458 AA.

AC Q99952; 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PROTEIN-TYROSINE-PHOSPHATASE (EC 3.1.3.48).
GN BDPL.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE=BRAIN;
RA Kim Y.W., Wang H.Y., Sures I., Lammers R., Martell K.J., Ullrich A.;
RX MEDLINE=97108574; PubMed=8950995;
RT "Characterization of the PST family protein tyrosine phosphatase
BDPL";
RL Oncogene 13:2275-2279(1996).
DR EMBL; X79568; CAA56105.1; -;
DR HSP; Q6124; 2SHP.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00700; PRTYPHPTASE.
DR SMART; SM00194; PTPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolyase.

QY 25 ANEFLKLRQSKYKADKTYPTTVAENAKNKKYKIDLPYDYSRVELSLTSDSSY 84
Db 27 ACEFSDIQACSAWKADGVCSTVAGSRPENVRKNRYKDVLPYDQTRVILSLQEGHSDY 86
QY 85 INANIKGYGPKAYIATQGLSTLLDFWRMIWEYSVLIIWACMEYEMGKKCERYWA 144
Db 87 INGNFIRGVDGLAYIATQGLPHTLLDFWRLVWEEFGVKVILMACREIENGRKRCERYWA 146
QY 145 EPEGMQLFGPFSVSCAEK-RKSDYIIRTLKVKFNSETRTIYQHYKNWPDHVPSSTD 203
Db 147 QQQE-PLQTGLFCITLIEKWLNEIMLTKLVTKQKRSRYVQLQWSWPDGVPSSPD 205
QY 204 PILELIMVRCYQEDSDVPICHCAGCGRTGCAIYDVTWMLKDGIIPIENFSVSLI 263
Db 206 HMLAMVEEARLQGSPEPLCVHCAGCGRTGVLCT-VDYVRQLLTQMIPDFSLFDV 264
QY 264 REMRQPSLVTQGYELVYNAVLELFRQMDVIRDKHSGTSQAKHCIP 314
Db 265 LKMRQRPAAVQTEQYRFLYHTVAQMF---CSTLQNASPHYQNIKENCAP 312

Query Match 19.6%; Score 708; DB 4; Length 458;
Best Local Similarity 46.7%; Pred. No. 2.2e-42;
Matches 136; Conservative 54; Mismatches 99; Indels 6; Gaps 4;

RESULT 10
Q9H0Y6 ID Q9H0Y6 PRELIMINARY; PRT; 1114 AA.

AC Q9H0Y6; 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE T200 LEUKOCYTE COMMON ANTIGEN (FRAGMENT).
GN PTPRC.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Ralph S.J., Thomas M.L., Morton C.C., Trowbridge I.S.;
RX MEDLINE=87275816; PubMed=2956090;
RT "Structural variants of human T200 glycoprotein (leukocyte-common
antigen)";
RL EMBO J. 6:1251-1257(1987).
RN [2]
RP SEQUENCE FROM N.A.

RC TISSUE=PLACENTA;
RX MEDLINE=89009812; PubMed=2971730;
RA Hall L.R., Streuli M., Schlossman S.F., Saito H.;
RT "Complete exon-intron organization of the human leukocyte common
antigen (CD45) gene";
RL J. Immunol. 141:2781-2787(1988).
DR EMBL; M23492; AAD15273.2; -;
DR EMBL; M23496; AAD15273.2; JOINED.
DR EMBL; M23466; AAD15273.2; JOINED.
DR EMBL; M23467; AAD15273.2; JOINED.
DR EMBL; M23468; AAD15273.2; JOINED.
DR EMBL; M23469; AAD15273.2; JOINED.
DR EMBL; M23470; AAD15273.2; JOINED.
DR EMBL; M23471; AAD15273.2; JOINED.
DR EMBL; M23472; AAD15273.2; JOINED.
DR EMBL; M23473; AAD15273.2; JOINED.
DR EMBL; M23474; AAD15273.2; JOINED.
DR EMBL; M23475; AAD15273.2; JOINED.
DR EMBL; M23476; AAD15273.2; JOINED.
DR EMBL; M23477; AAD15273.2; JOINED.
DR EMBL; M23478; AAD15273.2; JOINED.
DR EMBL; M23479; AAD15273.2; JOINED.
DR EMBL; M23480; AAD15273.2; JOINED.
DR EMBL; M23481; AAD15273.2; JOINED.
DR EMBL; M23482; AAD15273.2; JOINED.
DR EMBL; M23483; AAD15273.2; JOINED.
DR EMBL; M23484; AAD15273.2; JOINED.
DR EMBL; M23485; AAD15273.2; JOINED.
DR EMBL; M23486; AAD15273.2; JOINED.
DR EMBL; M23487; AAD15273.2; JOINED.
DR EMBL; M23488; AAD15273.2; JOINED.
DR EMBL; M23489; AAD15273.2; JOINED.
DR EMBL; M23490; AAD15273.2; JOINED.
DR EMBL; M23491; AAD15273.2; JOINED.
DR HSP; P18052; IYPO.

DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003595; PTPC_motif.
DR InterPro; IPR000387; TYR_phosphatase.
DR InterPro; IPR000242; Tyr_prot_phptase.
DR Pfam; PF00041; fn3; 2.
DR PRINTS; PR00700; PRTYPHPTASE.
DR SMART; SM00060; FN3; 2.
DR SMART; SM00194; PTPC; 2.
DR SMART; SM00404; PTPC_motif; 2.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 2.
DR PROSITE; PS00055; TYR_PHOSPHATASE_PTP; 2.
KW Hydrolyase.

FT NON_TER
SQ SEQUENCE 1114 AA; 127625 MW; 055385D916234CC8 CRC64;

Query Match 14.3%; Score 515.5; DB 4; Length 1114;
Best Local Similarity 28.4%; Pred. NO. 3.8e-28;
Matches 167; Conservative 87; Mismatches 238; Indels 97; Gaps 20;

QY 2 DOREILQK-----FLDAQSKKITKEE--FANEFLKLRQSKYKADKTY 44
Db 422 EQEELVERDEKOLMNVPEPIHADILLETYYKRIADGRLFLAEFOSIPRVFSK-----F 475

SQ SEQUENCE 1143 AA: 130897 MW; DB4B4400F3602B3C CRC64;

 Query Match 14.3%; Score 515.5; DB 4; Length 1143;
 Best Local Similarity 28.4%; Pred. No. 3.9e-28;
 Matches 167; Conservative 87; Mismatches 238; Indels 97; Gaps

 QY 2 DQREILQK-----FLDAQSKKITKEE--FANEFLKLKRSTKYKADKTY 44
 Db :|::|::| |::|::| |::|::| |::|::| |::|::| |::|::| |:
 451 EQQLVERDEKQLNMVEPIHADILLETYYKKRIADEGRFLAEPQSIPRVFSK-----F 504

 QY 45 PTTVAENAKNKKRYXDILPYDSRVELSLIISDESSVINANFIKVYGPAYATQG 104
 Db |::|::| |::|::| |::|::| |::|::| |::|::| |::|::| |:
 505 PIKEARPFQNKNRYDILLPYDYNRYVELSEINGDAGSNINASYIDGFKEPKRYIAAQG 564

 QY 105 PLSTTLDFWRMIWEYSVLVIVMACMEYEMGKKRCERYWAEPEGMQLFEGFPFVSCEAEK 164
 Db |::|::| |::|::| |::|::| |::|::| |::|::| |::|::| |:
 565 PRDETVDDFWRMIWEQKATVIVMWTRCEEGRNRNKCAEYWFPMEEGTAFGDVVVKINQHK 624

 QY 165 RKS DYIIIRLTUV----KFNSERTTIYQPHYKNPDDHVPSSIDPILELIMDVRCYQEDDSV 221
 Db |::|::| |::|::| |::|::| |::|::| |::|::| |::|::| |:
 625 RCPDYIIOKLINVNKKEKATGREVTTHIQFTSWPDHGYPEDPHLLLKLRRRVNAFSPFSG 684

 QY 222 PICIHCSAGCGRTGVICAIVDYTWMLKDGIIPBN-FSVFSLIREMTQRPSLVQTGEQY 280
 Db ||::|::| |::|::| |::|::| |::|::| |::|::| |::|::| |:
 685 PIVVHC SAGVGRTGYTIGI-----DAMLEGLEAENKVDVYGYYVKLRQRCLMAQVVEAY 739

 QY 281 ELVYNVAVLEL-FKRQMDVIRDKHSGTESQAKHCIPEKNHTLOADSYSPNLPR-STTKAA 337
 Db ||::|::| |::|::| |::|::| |::|::| |::|::| |::|::| |:
 740 ILIHOALVEYNQGETEVNISELHPYLHNKKRDPPSEPSPLEAE--FORLPYSYRSWRTO 797

 QY 338 KMMNQORTKWEIKESSDFRTSISAKEELVL-----HPAKSTSFDFLELAYSFDKNA 392
 Db ||::|::| |::|::| |::|::| |::|::| |::|::| |::|::| |:
 798 HIGQNEENKSKNRNSVIPYDYNRPVPLKHELEMSKESEHDSDSDSDSEEPSKYINA 857

 QY 393 DTTWK-WOTKAPPIVGEPLQKHOSLDGLSLF-----EGCSNSKPVN 433
 Db ||::|::| |::|::| |::|::| |::|::| |::|::| |::|::| |:
 858 SFINSYKWPYMIAAAGPL-KETIGDFQMIFORKXVIVMLTELKHGDOEICAQ---YW 913

 QY 434 AAGR--FYNSKVPITRT-KSTPEEL--IQORENKEVDKSNESY-----LESOPHD 479
 Db |:|::|::| |::|::| |::|::| |::|::| |::|::| |::|::| |:
 914 GEGRQTYTGDI EVDUKDTKSTYTLYRFVELRHSKRKDRTRYQTYQYNWSVEQLPAEPKE 973

 QY 480 -SCFVENQAQKMHVSSAELNISLPYDSKKHQIRNASNVKHHDSALGVY 527
 Db ::||::|::| |::|::| |::|::| |::|::| |::|::| |::|::| |:
 974 LISMIQVVKQLPKQNSSEGN-----KHKHSTPLLIIH 1005

 RESULT 12
 Q61373 PRELIMINARY; PRT; 361 AA.
 AC Q61373;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE PROTEIN TYROSINE PHOSPHATASE (FRAGMENT).
 GN PTPRJ OR PTP-RL9.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57/BL6; TISSUE=LIVER;
 RA Hagahitsuji H., Arii S., Furutani M., Imamura M., Kaneko Y.,
 RA Takenawa J., Nakayama H., Fujita J.;
 RT "Enhanced expression of multiple protein tyrosine phosphatases in the
 RT regenerating mouse liver : Isolation of ptp-rl 10, a novel
 RT cytoplasmic-type phosphatase with sequence homology to cytoskeletal
 RT protein 4.1.";
 RL Oncogene 0:0-0(1995).
 DR EMBL; D49393; BAA08386.1; -.
 DR HSSP; P18052; LYFO.


```

DR MGD; MGI:104574; Ptpn1.
DR InterPro; IPR000387; Tyr_phosphatase.
DR Pfam; PF000242; Tyr_prot_phptase.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00020; MAMDOMAIN.
DR PRINTS; PR00700; PRTYPHPTASE.
DR SMART; SM00194; PTPc; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS0056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS0055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase.
FT NON_TER 1
SQ SEQUENCE 361 AA; 41726 MW; B51462E8D11023AA CRC64;

Query Match 14.2%; Score 513; DB 11; Length 361;
Best Local Similarity 39.08; Pred. No. 1.3e-28;
Matches 115; Conservative 50; Mismatches 110; Indels 20; Gaps 7;

Qy 15 QSKKITKEEFANFLKLRQST-----KYKADK----TYPTTVAENAKNKNRYKDILP 65
Db 41 KSKLIRVENFEAYFKKQQAQNSCGFAEYEDKLIGISLPKYTAETAEIENRCKNRYNNVLP 100

Qy 66 YDSRVLSLTSDESSYINANIKYGVKAYATQGPLSTLLDFWRMIWEYSVLII 125
Db 101 YDISRVLSVQTHSTD-DYINANTMPGVHSHKDFIATQGPLNTLKDQFWRMVWEKNYAI 159

Qy 126 VMACEYEMGKKCBRYWAEPGEMQLEFPFVSCEAEKRSKYIIRTLKVK--FNSETR 183
Db 160 VMLTKCVSQGRKCEEW--PSKAQDYGDIIVAMTSEVLPEWTIRDFVYKNNQNSSEH 217

Qy 184 TIYQHYKNWPDHDPSSIDPILELINDVRCY--QEDDSVPICIHCSAGCGTGVCIAIV 241
Db 218 PLRQFHETSPDGHGVPDPTDLLINFRYLVRDYMKQIPPEPILVHCSAGVGRGTGTFIAD 277

Qy 242 DYTWMLLKDGILPENFSVSLIREMRTORPSLVQTOEQYELVYNVLELFRQMD 296
Db 278 RLIIQIENETV---DVGIVYDLMRHRLMWQTDQYVFLNQCVDLDIRAQKD 328

RESULT 13
Q9PUM2 PRELIMINARY; PRT; 1195 AA.
AC Q9PUM2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE RECEPTOR PROTEIN TYROSINE PHOSPHATASE RHO PRECURSOR (FRAGMENT).
GN PTPRT.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20193505; PubMed=10727868;
RA Johnson K.G., Holt C.E.;
RT "Expression of CRYP-alpha, LAR, PTP-delta, and PTP-rho in the
RL developing xenopus visual system.";
RL Mech. Dev. 92:291-294(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Besco J.A., Johnson K.G., Frostholt A., Popesco M., Burghes A.H.M.,
RA Holt C.E., Rottler A.;
RT "Xenopus laevis receptor protein tyrosine phosphatase rho.";
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 MAM DOMAIN.
DR EMBL; AF173857; AAD50295.1; -.
DR HSSP; P28827; 1RPM.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR000998; MAM.
DR InterPro; IPR000387; Tyr_phosphatase.
DR InterPro; IPR000242; Tyr_prot_phptase.

```

```

DR Pfam; PF00041; fn3; 3.
DR Pfam; PF00629; MAM; 1.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00020; MAMDOMAIN.
DR PRINTS; PR00700; PRTYPHPTASE.
DR SMART; SM00060; FN3; 3.
DR SMART; SM00137; MAM; 1.
DR SMART; SM00194; PTPc; 1.
DR PROSITE; PS00740; MAM_1; UNKNOWN_1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS0056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS0055; TYR_PHOSPHATASE_PTP; 1.
KW Glycoprotein; Hydrolase; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 >1195 RECEPTOR PROTEIN TYROSINE PHOSPHATASE
FT RHO.
FT NON_TER 1195 1195
SQ SEQUENCE 1195 AA; 132818 MW; 71F3C6D0AAD34F07 CRC64;

Query Match 14.1%; Score 511; DB 13; Length 1195;
Best Local Similarity 41.58; Pred. No. 8.7e-28;
Matches 107; Conservative 47; Mismatches 92; Indels 12; Gaps 4;

Qy 37 KYKA---DKTYPTTVAENAKNKNRYKDILPYDSRVLSLTSDESSYINANIKGV 93
Db 908 EYEALPEQPTASDWTAKEDENRNKNRYGNIISYDHSRVLRQIEEDPHSDYINANYVDGY 967

Qy 94 YGPKAYATQGPLSTLLDFWRMIWEYSVLIIWACMEYEMGKKCBRYWAEPGEMQLEF 153
Db 968 HRPRYIATQGPQMETVKDFWRMIWQENSISVVMVTNLVEGRVKCVRYWPDDEV---Y 1024

Qy 154 GPFVSCEAEKRSKYIIRTLKVK--KENSEPTIYQHYKNWPDHDPSSIDPILELMD 211
Db 1025 GDIKVAIETEALEYIIRTVQKGSHEIREIRQHFHTSWPDNGVPCYATGLGFIRQ 1084

Qy 212 VRCYQEDDSVPICIHCSAGCGTGVCIAIVDTYWMLLKDGILPENFSVSLIREMRTORP 271
Db 1085 VKFLNPPDAGPITVHCSAGAGRTGCFIADIMLDMAEQGV---DIFNCVRELRAQRV 1140

Qy 272 SLVQTOEQYELVYNVLE 289
Db 1141 NMVQTEQYVFEHDAILE 1158

RESULT 14
Q62884 PRELIMINARY; PRT; 1216 AA.
AC Q62884;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE DENSITY-ENHANCED PHOSPHATASE-1 PRECURSOR (EC 3.1.3.48) (DEP-1)
DE (VASCULAR PROTEIN TYROSINE PHOSPHATASE 1).
GN DEP1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR KYOTO; TISSUE=AORTA;
RX MEDLINE-96375109; PubMed=8781490;
RA Borges L.G., Selfert R.A., Grant F.J., Hart C.E., Distèche C.M.,
RA Edeihoff S., Solca F.F., Lieberman M.A., Lindner V., Fischer E.H.,
RA Lok S., Bowen-Pope D.F.;
RT "Cloning and characterization of rat density-enhanced phosphatase-1, a
RL protein tyrosine phosphatase expressed by vascular cells.";
RL Circ. Res. 79:570-580(1996).
CC -!- FUNCTION: MAY PLAY A ROLE IN ADHESION AND/OR SIGNALING EVENTS
CC INVOLVING CELL-CELL AND CELL-MATRIX CONTACT.
CC -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O = PROTEIN

```


QY 66 YDYSRVELSLITSDSSYINANFIKGYGPKAYIATQGPLSTTLTLLDFWRMIWEYSVLII 125
Db 1077 YDISRVKLSVQTHSTD-DYINANYMPGVHKKDFIATQGPLNTLKDFRWWKKNVYAI 1135
QY 126 VMACMEYEMGKKCERYNAEPGEMOLEFGPFSVSCAEKRXSDYIIRTLKVK--FNSETR 183
Db 1136 IMLTKCVEQGRTKCEEYW--PSKQADYGDITVAMTSEIVLPFWTIRDTVKNIQTSSEH 1193
QY 184 TIYQFHYKNWPDHVPSSIDPILWIDVRCY--QEDDSVPICIHCSAGCGRTGVICAIV 241
Db 1194 PLRQFHFTSWPDHGVDPDTDLLINFRYLVRDYMKOSPESPILVHCSAGVGRGTGTFIAID 1253
QY 242 DYTWMLLKDGIIPEFVSFSLIREMRTORPSLVQTEQYELVYNNAVLELEFRQMD 296
Db 1254 RLIIYQIENENTV----DVYGIYVDLRMHRLMVQTEQYVFLNQCVDIVRSQKD 1304

Search completed: May 26, 2002, 12:26:01
Job time: 546 sec

OM of: US-09-600-358a-4 to: EST:* out_format : ffs

Date: May 26, 2002 1:04 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

```
-MODEL=framet p2n.model -DEV=xlp  
-O=/cpn2.1/USPTO_spool/US09600358/runat_24052002_101553_26989/app_query.fasta_1.761  
-DB=EST -QFMT=fastap -SUFFIX=1st -GAPOP=12.000 -GAPEXT=4.000  
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500  
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000  
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000  
-DELEXT=7.000 -START=1 -MATRIX=blomsm62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pts -NORM=ext -HEADSIZE=500  
-MINLEN=0 -MAXLEN=200000000 -USER=US09600358.@CGN1_17299  
-ICPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30  
-NO_XLPXY -WAIT -THREADS=1
```

Search information block:

Query: US-09-600-358a-4

Query length: 692

Database: EST*

Database sequences: 13736207

Database length: -1841457050

Search time (sec): 2277.120000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_hic:AF077031	+	3512.50	5975.82	0.0	2903	! AF077031 Homo sapiens protein
gb_est2:BM458477	+	1578.00	2675.89	8.4e-140	1116	! BM458477 AGENCOURT 6413950 NIH
gb_est2:BG541535	+	1112.00	1881.73	1.4e-95	832	! BG541535 602570982F1 NIH_MGC_77
gb_est1:AU132052	+	929.50	1569.87	3.4e-78	796	! AU132052 AU132052 NT2RP3 Homo s
gb_est2:BM469079	+	887.00	1494.34	5.5e-74	1003	! BM469079 AGENCOURT 6445886 NIH
gb_est1:AW967600	+	868.50	1469.21	1.4e-72	582	! AW967600 EST379675 MAGE resequ
gb_est1:AW504119	+	848.50	1437.42	8.1e-71	475	! AW504119 UI-HF-BN0-all-h-10-0-0
gb_est1:AU132040	+	835.00	1407.88	3.6e-69	811	! AU132040 AU132040 NT2RP3 Homo s
gb_est1:AW373463	+	824.00	1394.37	2.0e-68	521	! AW373463 RC0-BT0522-071299-011
gb_est1:AW363814	+	788.00	1331.29	6.6e-65	588	! AW363814 RC0-CT0358-071299-011
gb_est1:AW915666	+	760.00	1284.08	2.8e-62	554	! AW915666 EST346970 Normalized x
gb_est2:BG498752	+	738.00	1242.86	5.6e-60	745	! BG498752 602544381F1 NIH_MGC_60
gb_est2:BL152499	+	718.50	1208.11	4.8e-58	835	! BL152499 602915640F1 NCI_CGAP_1
gb_est1:AW822729	+	694.50	1170.43	6.0e-56	629	! AW822729 uq17g09.y1 Ren Scubbs
gb_est2:BM194577	+	661.50	1116.58	6.0e-53	505	! BM194577 TCAAP3E10089 Pediatr
gb_est2:BI411856	+	646.00	1081.20	5.6e-51	1054	! BI411856 602966483F1 NCI_CGAP
gb_est1:AJ394015	+	621.50	1041.72	8.9e-49	859	! AJ394015 AJ394015 dkf426 Gallu
gb_est1:AL597518	+	611.00	1031.58	3.3e-48	448	! AL597518 DKZP313F0814_r1 313 C
gb_est2:BF301903	+	571.00	954.30	6.6e-44	932	! BF301903 602032931F1 NCI_CGAP_5
gb_est1:AA851710	+	569.00	962.48	2.3e-44	355	! AA851710 EST194478 Normalized x
gb_est2:BF105998	+	561.00	938.94	4.7e-43	805	! BF105998 601823023F1 NIH_MGC_77
gb_est1:AI530640	+	556.00	930.78	1.3e-42	836	! AI530640 mp70q01.y1 Soares thym
gb_est2:BE765069	+	550.00	925.84	2.5e-42	500	! BE765069 CMI-NT0087-280600-291
gb_est2:BG614041	+	545.00	911.42	1.6e-41	814	! BG614041 602641583F1 NIH_MGC_61
gb_est2:BG505436	+	539.00	904.42	3.9e-41	620	! BG505436 602549331F1 NIH_MGC_61
gb_est2:BI415856	+	533.50	891.72	2.0e-40	815	! BI415856 602988052F1 NCI_CGAP_1
gb_est1:AA998917	+	530.00	896.60	1.1e-40	330	! AA998917 UI-R-C-02-0-UI-s
gb_est1:AL522809	+	508.50	851.10	3.7e-38	680	! AL522809 AL522809 LTI_NFL004.NE
gb_est1:AL600035	+	502.00	845.19	7.9e-38	441	! AL600035 DKZP3130203_r1 313 C
gb_est2:BI691928	+	487.00	830.56	5.1e-37	730	! BI691928 603511861F1 NCI_CGAP_M
gb_est1:AU121737	+	485.00	810.16	7.0e-36	722	! AU121737 AU121737 NT2RP2 Homo s
gb_est2:BF099411	+	485.00	807.33	1.0e-35	913	! BF099411 601751512F1 NCI_CGAP_M
gb_est1:AI323259	+	481.50	808.35	8.9e-36	510	! AI323259 mp91b04.y1 Soares thym
gb_est1:BI415322	+	479.50	795.79	4.4e-35	1090	! BI415322 602987967F1 NCI_CGAP
gb_est1:AA241226	+	477.00	798.47	3.1e-35	611	! AA241226 mu95d08.r1 Soares mous
gb_est2:BF579759	+	476.00	791.09	8.1e-35	979	! BF579759 602095070F1 NCI_CGAP_C
gb_est2:BG694807	+	472.50	788.88	1.1e-34	715	! BG694807 602829360F1 NCI_CGAP_C
gb_est2:BE765067	+	455.00	765.29	2.2e-33	421	! BE765067 CMI-NT0087-280600-291
gb_est1:AI981941	+	454.00	759.83	4.5e-33	575	! AI981941 pat.pk0069.d6.f chicken
gb_est2:BI686558	+	451.00	749.41	1.7e-32	892	! BI686558 603313654F1 NCI_CGAP_M

gb_est2:BF144494 + 448.50 747.71 2.1e-32 720 ! BF144494 601790135F1 NCI_CGA
gb_est2:BG234079 + 447.50 754.14 9.3e-33 366 ! BG234079 daa36e12.y1 NICH
gb_est1:AW766206 + 447.00 746.13 2.6e-32 663 ! AW766206 da83f09.y1 Harlan
gb_est1:AL555127 + 446.00 743.40 3.7e-32 722 ! AL555127 AL555127 LTI_NFL006
gb_est1:AW319424 + 442.00 737.20 8.1e-32 684 ! AW319424 un16c07.y1 Sugano m

seq_name: gb_hic:AF077031

seq_documentation_block:

LOCUS AF077031 2903 bp mRNA linear HTC 22-MAY-2001
DEFINITION Homo sapiens protein tyrosine phosphatase homolog mRNA, complete cds.

ACCESSION AF077031

VERSION AF077031.1 GI:4689109

KEYWORDS HTC.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2903)

AUTHORS Liu,T., Zhang,J., Fu,G., Zhang,Q., Ye,M., Zhou,J., Wu,J., Shen,Y., Yu,M., Chen,S., Mao,M. and Chen,Z.

TITLE Human protein tyrosine phosphatase (70zpep) homolog

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2903)

AUTHORS Liu,T.

TITLE Direct Submission

JOURNAL

Submitted (10-JUL-1998) Shanghai Institute of Hematology, Shanghai
Second Medical University, Rui-Jin Hospital, 197 Rui-Jin Road II,
Shanghai 200025, P.R. China

FEATURES

Location/Qualifiers

1..2903

/organism="Homo sapiens"

/db_xref="taxon:9606"

/cell_line="NB4"

73..2496

/codon_start=1

/product="protein tyrosine phosphatase homolog"

/protein_id="AAD27764.1"

/db_xref="GI:4689110"

/translation="MDREILQFLDEAQSQTKEEFANEFKLKLRSTKYKADKTY
PTAAETKPKNIKKRKYKIDLPYDYSRVELSLTSDSSYINANFKYVGPAYAT
STAEKTLDFWRMTWEXSVLIIGMACMEYKWKCCRYWAEVLEFGPFSVSC
EAEKRSKYIITLAVKNSRTTYIQHYKWNPDHVPSSIDPILWVRCYQED
DSVPICHCACGRTGICADITWMLLKDGIIPENSYSFLIREMTQRSLVQTQ
EYELVYNALVELFRQMDIKRHSGETSQAKHCIPKRNTHLQADSYSPNLPKSTTK
AARMNQRTKMEIKESSDFRTSEISAKELVLHPAKSYDFLELNYSPDKNAD
TTMKWQTAFFPIVGEPLQKHSLDGLSPLEFGCSNKPVNAAGRYFNKVPITRTKST
PELLOQRETKVDSEKFNFSYLESOPHDSCEFVEMQAKVMHVSSAELNYSLPYDSKHQ
IRNASNVKHDSALGVSYIPLVNPYESSWPPSGTSKMSLDLPEKODGVFPFSSL
LPTSTSLFSYNSHDSLSLSNPTNISLLNSESALVATAPRIDDEIPPLPFWTPES
FIVVEAGFEFSPNPKSSASAVKVGIGTSLMGWTSSEPKKFDSDVILRPSKVLKRP
KSELQDRSSPPPLPRTLEFSLADEDCMQAQSIETYSYSPDTMENSTSSKQTLK
TSKGFTKTSKSLILRNMKKSCNCPNPAESVQSNSSSFLNFGANRFSKPKGP
RNPPTWNI"

BASE COUNT 955 a 580 c 528 g 840 t

ORIGIN

alignment_scores:

Quality: 3512.50 Length: 685
Ratio: 5.173 Gaps: 1
Percent Similarity: 99.124 Percent Identity: 98.832

alignment_block:

US-09-600-358A-4 x AF077031 ..

Align seg 1/1 to: AF077031 from: 1 to: 2903

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17
|||||
73 ATGACCAAGAGAAATTCGACAGAGTCTCTGGATGAGGCCCAAGCAA 122

17 slyslleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArgG 34
|||||
123 GAAATTTACTAAAGAGGAGTTTGCCAAATGAATTTCTGAAGCTGAAAGGC 172

34 lnSerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50
|||||
173 AATCTACCAAGTACAAGGAGAGACAAACCTATCTCTACAACTGTGGCTGAG 222

51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProThrAs 67
::: |||||
223 AAGCCCAAGAAATATCAAGAAACACAGATATAGGATATTTTGGCCCTATGA 272

67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSerT 84
|||||
273 TTATAGCCGGTAGAACTATCCCTGATACCTCTGATGAGGATTCAGCT 322

84 yrIleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
|||||
323 ACATCAATGCCAACTTCATTAAAGGAGTTTATGGACCCAGGCTTATATT 372

101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetil 117
|||||
373 GCCACCCAGGGCTCTTATCTACAACCCCTCTGGACTTCTGGAGGATGAT 422

117 eTrpGluTyrSerValLeuIleIleValMetAlaCysMetGluTyrGluM 134
|||||
423 TTGGGAATATAGTCTCTTATCATTTGGTATGGCATGCATGGAGTATGAAA 472

134 etGlyLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
|||||
473 TGGGAAGAAAAAGTGTGAGCGCTACTGGGCTGAGCCAGTAGAGATGCCAG 522

151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
|||||
523 CTGGAATTTGGCCCTTCTCTGTATCTCTGTGAAGCTGAAAAAGGAAATC 572

167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
|||||
573 TGATATATAATCAGGACTCTAAAAGTTAAAGTTCAATAGTGAACATCGAA 622

184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisaspValProSer 200
|||||
623 CTATCTACCAAGTTTCATTACAAGAAATGGCCAGACCATGATGTACCTTCA 672

201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCystyrGlnGl 217
|||||
673 TCTATAGACCTTATCTTGAGCTCATCTGGGATGTACGTTGTTACCAAGA 722

217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
|||||
723 GGATGACAGTGTCCCATATGCATTCACATGCAGTGTGGCTGTGGGAAGGA 772

234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
|||||
773 CTGGTGTATTGTGCTATT...GATTATACATGGATGTGTCTAAAAGAT 819

251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetar 267
|||||
820 GGGATAAATTCCTGAGAACITTCAGTGTITTCAGTTTGAICGGGAAATGCG 869

267 gThrGlnArgProSerLeuValGlnThrGlnGluInTyrGluLeuValT 284
|||||
870 GACACAGAGGCCCTTCATTAGTTCAACAGCGCAGGAACAATATGAACCTGGTCT 919

284 yrAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
|||||
920 ACAATGCTGTATTAGAACTATTTAAGACAGACAGATGGATGTTATCAGAGAT 969

301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
|||||
970 AACAATTTCTGGACAGAGATCAAGCAAGCATTTGTATTCCTCCTGAGAAAA 1019

317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
|||||

1020 TCACACTCTCCCAAGCAGACTCTATTCTCTTAATTTACCAAAAAGTACCA 1069

334 hrLysAlaAlaLysMetMetAsnGlnGlnArgThrLysMetGluIleLys 350
|||||
1070 CAAAAGCAGCAAAAATGATGAACCAACAAGGACAAAAATGGAATCAAAA 1119

351 GluSerSerSerPheAspPheArgThrSerGluIleSerAlaLysGluGl 367
|||||
1120 GAATCTTCTCTCTTGTGACTTTAGGACTTCTGAATAAGTGCAAAAGAGA 1169

367 uLeuValLeuHisProAlaLysSerThrSerPheAspPheLeuGluL 384
|||||
1170 GCTAGTTTTGCACCCTGCTAAATCAAGCACTCTCTTTTGACTTTCTGGAGC 1219

384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnThr 400
|||||
1220 TAAATTCACAGTTTGTACAAAAATGCTGACACAACCATGAATGGCAGACA 1269

401 LysAlaPheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAs 417
|||||
1270 AAGGCATTTCCAAATAGTTGGGAGCCTCTTCAGAGCATCAAAAGTTTGA 1319

417 pLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsnA 434
|||||
1320 TTTGGGCTCTCTCTTGTGTTGAGGATGTTCTAATCTCTAAACCTGTAATG 1369

434 laAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSer 450
|||||
1370 CAGCAGGAAGATATTTAATTCAAAAGTGCCAAATAACACGACGACCAATCA 1419

451 ThrProPheGluLeuIleGlnArgGluThrLysGluValAspSerLy 467
|||||
1420 ACTCTTTTGAATGTATACAGCAGAGAGAAACCAAGAGGTGGACAGCAA 1469

467 sGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheValG 484
|||||
1470 GGAATACTTTTCTTATTGGAATCTCAACACATGATCTCTGTGTTGTAG 1519

484 luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
|||||
1520 AGATGAGGCTCAAAAAGTAATGATGTTCTTCACGACAGACTGAATAT 1569

501 SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnVally 517
|||||
1570 TCACTGCCATATGACTCTAAACACCAAAATACGTAATGCCTCTAATCTAAA 1619

517 sHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValG 534
|||||
1620 GCACCATGACTCTAGTGCTCTTGGTGTAATTTCTTACATACCTTTTAGTG 1669

534 luAsnProTyrPheSerSerTyrProProSerGlyThrSerSerLysMet 550
|||||
1670 AAAATCCTTATTTTTCATGCGCTCCCAAGTGGTACCACTTCTTAAGATG 1719

551 SerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerLe 567
|||||
1720 TCTCTTGATTTACCTGAGAAGCAAGATGGAACGTGTTTTTCTCTCTCTCT 1769

567 uLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSerS 584
|||||
1770 GTTGCCAAACATCTCTACATCCCTCTCTCTTATTACAAATTCACATGATT 1819

584 erLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnGlu 600
|||||
1820 CTTTATCACTGAAATCTCCAACCAATATTTCTCTCACTATTGAACCAAGGAG 1869

601 SerAlaValLeuAlaThrAlaProArgIleAspAspGluIleProProPr 617
|||||
1870 TCAGCTGTACTAGCAACTGCTCCAAGATAGATGATGAAATCCCCCTCC 1919

617 oLeuProValArgThrProGluSerPheIleValValGluAlaGlyG 634
|||||

1920 ACTTCCTGTATGGACACCTGTAATCATTTATTTGTTGGTGAAGAGCTGGAG 1969
 634 luPheSerProAsnValProLysSerLeuSerAlaValLysValLys 650
 1970 AATTCTCACCAGAAATGTTCCCAAAATCCCTATTCCTCAGCTGTGAAGTAAAA 2019
 651 IleglyThrSerLeuGluTrpGlyClyThrSerGluProLysLysPheAs 667
 2020 ATTGGAACATCACTGGAATGGGTGGAACATCTGAACCAAGAAATTTGA 2069
 667 pAspSerValIleLeuArgProSerLysSerValLysLeuArgSerProL 684
 2070 TGACTCTGTGATACTTAGACCAAGCAAGAGTGTAAACCTCCGAAGTCCCTA 2119

684 ysSer 685
 2120 AATCA 2124

seq_name: gb_est2:BM458477

seq_documentation_block:
 LOCUS BM458477 1116 bp mRNA linear EST 05-FEB-2002
 DEFINITION AGENCOURT_6413950 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5497108
 5', mRNA sequence.

ACCESSION BM458477
 VERSION BM458477.1 GI:18507517
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1116)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Lou Staudt

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM12126 row: 1 column: 05

High quality sequence stop: 717.

Location/Qualifiers

FEATURES

source

1..1116

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5497108"

/clone_lib="NIH_MGC_85"

/tissue_type="lymphoma, cell line"

/lab_host="PH10B (phage-resistant)"

/notes="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 1.867 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

BASE COUNT 350 a 232 c 232 g 302 t

ORIGIN

alignment_scores:

Quality: 1578.00 Length: 404

Ratio: 4.767 Gaps: 3

Percent Similarity: 81.931 Percent Identity: 78.960

alignment_block:

US-09-600-358A-4 x BM458477 ..

Align seg 1/1 to: BM458477 from: 1 to: 1116

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17
 75 ATGGACCAAGAGAAATTCGACAGAGTTCCTGCATGAGGCCCAAGCAA 124
 17 sLysIleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArg 34
 125 GAAATTAATAAGAGGAGTTTGCAATGAATTTCTGAAGCTGAAAGGC 174
 34 InSerThrLysThrLysAlaAspLysThrTyrProThrThrValAlaGlu 50
 175 AATCTACCAAGTACAGGACAGACAAACCTATCTACACACTGTGGCTGAG 224
 51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
 225 AAGCCCAAGAAATATCAAGAAAAACAGATATAAGGATATTTTGCCTATGA 274
 67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSert 84
 275 TTATAGCCGGTAGAAGTATCCCTGATAACCTCTGATGAGGATTCACAGCT 324
 84 yTrIleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
 325 ACATCAATGCCAATTCATTAAAGGAGTTTATGACCAAGGCTTATATT 374
 101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheThrArgMetIl 117
 375 GCCACCCAGGGTCTCTTATCTACAACCTCTGGACITCTCGAGGATGAT 424
 117 eTrpGluTyrSerValLeuIleIleValMetAlaCysMetGluTyrGluM 134
 425 TTGGGAATATAGTCTCCTTATCATTTGTTATGGCATGATGAGTATGAAA 474
 134 etGlyLysLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
 475 TGGAAAGAAAAAGTGGAGCGCTACTGGGCTGAGCCAGAGAGATGAG 524
 151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
 525 CTGGAATTTGGCCCTTCTCTGTATCTGTGAAGCTGAAAAAGGAATC 574
 167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
 575 TGATTATATAATCAGGACTCTAAAAGTTAAGTTCAATAGTGAATCGAA 624
 184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSer 200
 625 CTATCTACCAAGTTTCAATACAGAAATTTGGCCAGACCATGATGACCTTCA 674
 201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnGl 217
 675 TCTATAGACCTATCTTGAGCTCATCTGGGATGATGCTTGTACCAAGA 724
 217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
 725 GGATGACAGTGTCTCCATATGCACTGCTGAGTGTGGTGTGGAGGA 774
 234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
 775 CTGGTGTATTGTGCTATT...GATTATACATGATGTTGTCTAAAGAG 821
 251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267
 822 GGG..... 824
 267 gThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValT 284
 824 824
 284 yAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
 824 824
 301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317

825AGTCAAGCCAAAGCATTGTATTCTCGAGAAAA 856
317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
857 TAC.ACTCCTCAAGCAGACTTTTCTCTTAATTTACCAAAAGTACCC 905
334 hrLysAlaAlaLysMetMetAsnGlnArgThrLysMetGluLys 350
906 CAAAGCCGCAAAATTTGATGAACCAACCAAGGACAAATGGAAATCAA 955
351 GluSerSerSer.PheAspPheArgThrSerGluLysSerAlaLysGluG 367
956 GAATCTCTTCCCTTTGACTTTTAGACTCTCTGAATAGG.GCCAAAGAAG 1004
367 luleuValLeuHsProAlaLysSerSerThrSerPheAspPheLeuGlu 383
1005 AGTAGTTTTCACCTGCTTAATCAGGACTTTCTTTGACTTTCGGGA 1054
384 LeuAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnTh 400
1055 GCAAAATTACGTTTGTGAAAATGCTGAACCAACCTTGAAATGGCCGAA 1104
400 rLysAlaPhe 403
1105 CAAGGCATTT 1114

seq_name: gb_est2:BG541535

seq_documentation_block:
LOCUS BG541535 832 bp mRNA linear EST 03-APR-2001
DEFINITION 602570982P1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4695140 5',
mRNA sequence.

ACCESSION BG541535
VERSION BG541535.1 GI:13533768
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 832)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LLCM1520 row: d column: 21
High quality sequence stop: 785.

Location/Qualifiers

1..832

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4695140"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B (T1 phage-resistant)"

/note="organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgctcgcc); Site_2: SfiI (ggccattatgcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGGCATTATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGGCGGCGGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

BASE COUNT 245 a 198 c 136 g 252 t 1 others

ORIGIN

alignment_scores:
Quality: 1112.00 Length: 223
Ratio: 5.032 Gaps: 1
Percent Similarity: 99.103 Percent Identity: 98.206

alignment_block:
US-09-600-358A-4 x BG541535 ..

Align seg 1/1 to: BG541535 from: 1 to: 832

464 ValAspSerLysGluAsnPheSerTyrLeuGluSerGlnProHisAspSe 480
1 GTGGCAGCAAGAAACTTTCTTATTGGGAATCTCAACACCATGATTC 50
480 rCysPheValGluMetGlnAlaGlnLysValMetHisValSerSerAlag 497
51 TTGTTTTGTAGAGATCGAGGCTCAAAAAGTAATGCATGTTTCTTCAGCAG 100
497 luleuAsnTyrSerLeuProTyrAspSerLysHisGlnIleArgAsnAla 513
101 AACTGAATTTACTACTGCCATATGACTCTAAACACCAATACGTAATGCC 150
514 SerAsnValLysHisHisAspSerSerAlaLeuGlyValTyrSerTyrI1 530
151 TCTAATGTAAGCACCACCATGACTCTAGTGTCTTGGTGATATATTCTTACAT 200
530 eProLeuValGluAsnProTyrPheSerSerTrpProSerGlyThrS 547
201 ACCTTTAGTGAATAATCCCTATTATTTTCATCATGCGCTCCCAAGTGGTACCA 250
547 erSerLysMetSerLeuAspLeuProGluLysGlnAspGlyThrValPhe 563
251 GTTCTAAGATGCTCTGTGATTTACCTTGAGAGCAAGATGGAACTGTGTTTT 300
564 ProSerSerLeuLeuProThrSerSerThrSerLeuPheSerTyrTrAs 580
301 CCTTCTCTCTGTGGCAACATCCTCTACATCCCTCTCTCTATTACAA 350
580 nSerHisSerSerLeuSerLeuAsnSerProThrAsnIleSerSerLeuL 597
351 TTCACATGATTTCTTATCAGTGAATTTCTCAACCAATATTTCCTCACTAT 400
597 euAsnGlnGluSerAlaValLeuAlaThrAlaProArgIleAspaspGlu 613
401 TGAACCAAGGAGTCAGCTGTACTAGCAACTGCTCCCAAGGATAGATGA 450
614 IleProProProLeuProValArgThrProGluSerPheIleValValG1 630
451 ATCCCCCTCCACTTCTCTGACGGACACCTGAATCATTTATNGTGTGTA 500
630 uGluAlaGlyGluPheSerProAsnValProLysSerLeuSerSerAlav 647
501 GGAAGCTGGAGAATTTCTCACCATAATGTTCCCAATCTTTATCCTCAGCTG 550
647 allYsValLysIleGlyThrSerLeuGluTrpGlyGlyThrSerGluPro 663
551 TGAAGTAAAAATTTGGAACATCACTCTGGAATGGGGTGGAAACATCTGA 600
664 LysLysPheAspAspSerValIleLeuArgProSerLys.SerValLysL 680
601 AAGAATTTGATGACTCTGTGATACTTTAGACCAAGCAAGAGTGTAAAC 650
680 euArgSerProLysSer 685
651 TCCGAAGTCTCTAAATCA 667

seq_name: gb_est1:AU132052

seq_documentation_block:

LOCUS AU132052

796 bp mRNA linear EST 24-OCT-2000

```

DEFINITION AU132052 NT2RP3 Homo sapiens cDNA clone NT2RP3003698 5', mRNA
sequence.
ACCESSION AU132052
VERSION AU132052.1 GI:10992406
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 796)
AUTHORS Ota.T., Nishikawa.T., Suzuki.Y., Ishii.S., Saito.K., Kawai.Y.,
Yamamoto.J., Wakamatsu.A., Nakamura.Y., Nagai.T., Sugano.S. and
Isogai.T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
Location/Qualifiers
source
1..796
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RP3003698"
/clone_lib="NT2RP3"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/notes="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT 281 a 132 c 162 g 217 t 4 others
ORIGIN

alignment_scores:
Quality: 929.50 Length: 254
Ratio: 4.187 Gaps: 1
Percent Similarity: 87.402 Percent Identity: 66.535

alignment_block:
US-09-600-358A-4 x AU132052 ..

Align seg 1/1 to: AU132052 from: 1 to: 796

33 ArgGlnSerThrLysTyrLysAlaAspLysThrTyrProThrValAl 49
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 AGATTGCTACCAATATAGCAAGAAAGATATATCCACAGCCACTGG 50
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
49 aGluAsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProT 66
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51 AGAAAGACAGAAATGTTAAAGACACAGATACAGACATAGTCCAT 100
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
66 yRAspTyrSerArgValGluLeuSerLeuThrSerAspGluAspSer 82
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
101 TTGATCACAGCCGAGTTAAATTTGACATTAAGACTCTTCACAGATTCA 150
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
83 SerTyrIleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTy 99
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151 GACTATATCATGCAATTTTATAAGGGCGTCTATGGGCCAAAGCATA 200
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
99 rIleAlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgM 116
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
201 TGTAGCAACTCAAGGACCTTTAGCAATACAGTAATAGATTTTGGAGGA 250
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
116 etIleTrpGluTyrSerValLeuIleValMetAlaCysMetGluTyr 132
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
251 TGATATGGAGTATAATGTTGTGATCATTTGTAATGGCCTGCCGGAATTT 300

```

```

133 GluMetGlyLysLysLysCysGluArgTyrTirPAlaGluProGlyGlu 149
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
301 GAGATGGGAAGGAAAATATGAGCGCTATGGCTTTGTATGGAGAAGA 350
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
149 tGlnLeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgL 166
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| :
351 CCCCATACGCTTTGCACCATTTAAATTTCTGTGTGAGGATGAACAAGCA 400
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
166 ySerAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThr 182
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| :
401 GAACAGACTACTTCTCAGGACACTTACTTGAATTTCAAAATGAATCT 450
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
183 ArgThrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValPr 199
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
451 CGTAGGCTGTATCAGTTTCATTATGTAAGTGGCCAGACCATGATGTTCC 500
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
199 oSerSerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrG 216
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
501 TTCATCATTTGATTCTTCTGGACATGATAAGCTTAATGAGGAATATC 550
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
216 lnGluAspAspSerValProIleCysIleHisCysSerAlaGlyCysGly 232
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
551 AAGAACATGAGATGTTCTTATTTGTTATTCATTCAGTCAGTCGANGCTG 600
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
233 ArgThrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeu 249
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
601 AGAACAGTGCCATTTGTGCCATA...GATTATACGTGGGAATTTACTAAA 647
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
249 sAspGlyIleIleProGluAsnPheSerValPheSerLeuIleArgGlu 266
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
648 AGCTGGGAAAATCCAGAGGAATTTAATGTTATTTAATTAACAAGAAA 697
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
266 etArgThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeu 282
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
698 TGAGAACACAAAGCATTTCTGCTGTCNCAACAAAGGACCAATNTGA 747
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
283 ValTyrAsnAla 286
::: |||||
748 GGTCATAGAGCT 759

seq_name: gb_est2:BM469079
seq_documentation_block:
LOCUS BM469079 1003 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6445886 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5587352
5', mRNA sequence.
ACCESSION BM469079
VERSION BM469079.1 GI:18518121
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1003)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12356 row: 1 column: 09
High quality sequence stop: 728.
Location/Qualifiers
1..1003
/organism="Homo sapiens"
/db_xref="taxon:9606"

```

```
/clone="IMAGE:5587352"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dr primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT      331 a   196 c   203 g   269 t
ORIGIN
```

```
alignment_scores:
  Quality: 887.00      Length: 279
  Ratio: 3.857         Gaps: 4
  Percent Similarity: 82.437  Percent Identity: 60.215
```

```
alignment_block:
US-09-600-358A-4 x BM469079 ..
Align seg 1/1 to: BM469079 from: 1 to: 1003
```

```
63   ileLeuProTyAspTyrSerArgValGluLeuSerLeuIleThrSerAs 79
      ::::::::::::::::::::::::::::::::::::::::::::::
5   CTACTGCCATTGATCACAGCCGAGTTAAATTGACATTAAAGACTCCTTC 54

79   pGluAspSerSerTyrIleAsnAlaAsnPheIleIysGlyValTyrGlyP 96
      ::::::::::::::::::::::::::::::::::::::::::::::
55   ACAAGATTTCAGACTATATCAATGCAAAATTTTATAAAGGGCGCTATGGGC 104

96   roLysAlaTyrIleAlaThrGlnGlyProLeuSerThrThrLeuLeuAsp 112
      ::::::::::::::::::::::::::::::::::::::::::::::
105  CAAAACATATGTAGCACTCAAGGACCTTTAGCAAAATACAGTAATAGAT 154

113  pheTrpArgMetIleTrpGluTyrSerValLeuIleValMetAlaCy 129
      ::::::::::::::::::::::::::::::::::::::::::::::
155  TTTTGGAGGATGATATGGGAGTATAATGTGTGATCATTGTAATGCCCTG 204

129  smetGluTyrGluMetGlyLysLysLysCysGluArgTyrTrpAlaGluP 146
      ::::::::::::::::::::::::::::::::::::::::::::::
205  CCGAAGAAATTTGAGATGGGAAGGAAAAAATGTGAGCGCTATTGGCCTTGT 254

146  roGlyGluMetGlnLeuGluPheGlyProPheSerValSerCysGluAla 162
      ::::::::::::::::::::::::::::::::::::::::::::::
255  ATGGAGAGAGCCCATAAACGTTTGCACCATTTAAATTTCTGTGAGGAT 304

163  GluLysArgLysSerAspTyrIleIleArgThrLeuLysValLysPheAs 179
      ::::::::::::::::::::::::::::::::::::::::::::::
305  GAACAAGCAAGAACAGACTACTTCATCAGGACACTCTTACTTGAATTTCA 354

179  nSerGluThrArgThrIleTyrGlnPheHisTyrLysAsnTrpProAspH 196
      ::::::::::::::::::::::::::::::::::::::::::::::
355  AAATGAATCTCTGAGGCTGTATCAGTTTCATTATGTGAATGGCCAGACC 404

196  isAspValProSerSerIleAspProIleLeuGluLeuIleTrpAspVal 212
      ::::::::::::::::::::::::::::::::::::::::::::::
405  ATGATGTTCTTCATCATTTGATTTCTATCTCGACATGATAAGCTTAATG 454

213  ArgCysTyrGlnGluAspSerValProIleCysIleHisCysSerAl 229
      ::::::::::::::::::::::::::::::::::::::::::::::
455  AGGAAATATCAAGAACATGAAGATGTCCTATTGTTATTTCATTCAGTCG 504

229  aGlyCysGlyArgThrGlyValIleCysAlaIleValAspTyrThrTrpM 246
      ::::::::::::::::::::::::::::::::::::::::::::::
505  AGGCTGTGGGAAGAACAGGTGCCATTGTGCCATA...GATTATACGTGA 551

246  etLeuLeuLysAspGlyIleIleProGluAsnPheSerValPheSerLeu 262
      ::::::::::::::::::::::::::::::::::::::::::::::
552  ATTTACTTAAAGCTGGGAAATACAGAGGAATTAATGTATTAAATTA 601

263  ileArgGluMetArgThrGlnArgProSerLeuValGlnThrGlnGluGl 279
      ::::::::::::::::::::::::::::::::::::::::::::::
```

```
602  ATACAAGAAATGAGAACACAAAGGCGATTCTGCAGTACAAACAAGGAGCA 651

279  nTyrGluLeuValTyrAsnAlaValLeuGluLeuPheLysArgGlnMetA 296
      ::::::::::::::::::::::::::::::::::::::::::::::
652  ATATGAACCTGTTCATAGAGCTATTGCCCACTGTTTGAAAAACAGCTAC 701

296  spVal.....IleArgAsp.....LysHis 302
      ::::::::::::::::::::::::::::::::::::::::::::::
702  AACTATATCAAAATTCATGGAGCTCAGAAATTCCTGATGGAGTGAATGAA 751

303  SerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAsnHisTh 319
      ::::::::::::::::::::::::::::::::::::::::::::::
752  ATTAACACTGAAAACATGCTCAGCTCCATAGAGCCTGAAAAACAA..... 796

319  rLeuGlnAlaAspSerTyrSerProAsnLeuProLys 331
      ::::::::::::::::::::::::::::::::::::::::::::::
797  .....GATTCTCCTCCTCAAAANCAACCCNAGG 823
```

```
seq_name: gb_estl:AW967600
```

```
seq_documentation_block:
```

```
LOCUS      AW967600                      582 bp      mRNA      linear      EST 01-JUN-2000
DEFINITION EST379675 MAGE resequences, MAGJ Homo sapiens cDNA, mRNA sequence.
```

```
ACCESSION   AW967600
```

```
VERSION     AW967600.1  GI:8157437
```

```
KEYWORDS    EST.
```

```
SOURCE      human.
```

```
ORGANISM    Homo sapiens
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE   1 (bases 1 to 582)
```

```
AUTHORS     Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspari,R., Gay,C., Holt
            ,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
            Quackenbush,J.
```

```
TITLE       Assessment of gene expression patterns in a model of colon tumor
```

```
metastasis using a 19,200 element cDNA microarray
```

```
JOURNAL     Unpublished (2000)
```

```
COMMENT     Contact: John Quackenbush
```

```
The Institute for Genomic Research
```

```
9712 Medical Center Dr., Rockville, MD 20850, USA
```

```
Tel: 301 838 3528
```

```
Fax: 301 838 0208
```

```
Email: johnq@tigr.org
```

```
Plate: 243
```

```
Seq primer: Forward.
```

```
Location/Qualifiers
```

```
source
```

```
1..582
```

```
/organism="Homo sapiens"
```

```
/db_xref="taxon:9606"
```

```
/clone_lib="MAGE resequences, MAGJ"
```

```
/note="Vector: pBluescriptSKm"
```

```
BASE COUNT  164 a  137 c  101 g  180 t
```

```
ORIGIN
```

```
alignment_scores:
```

```
Quality: 868.50      Length: 182
```

```
Ratio: 4.879         Gaps: 1
```

```
Percent Similarity: 97.802  Percent Identity: 94.505
```

```
alignment_block:
```

```
US-09-600-358A-4 x AW967600 ..
```

```
Align seg 1/1 to: AW967600 from: 1 to: 582
```

```
502  LeuProTyAspSerLysHisGlnIleArgAsnAlaSerAsnValLysHI 518
      ::::::::::::::::::::::::::::::::::::::::::::::
3   TTGGCCCTCGAGGCCAAGAATTCGGCACGA...GGCTCTAATGTAAGCA 49
```

```
518  sHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValGluA 535
      ::::::::::::::::::::::::::::::::::::::::::::::
50  CCATGACTCTAGTGCCTTGGTGTATTCTTACATACCTTTAGTAGGAAA 99
```



```

535 snProTyPheSerSerTrpProProSerGlyThrSerSerLysMetSer 551
|||||
100 ATCTTATTTTTCATCATGGCTCCCAAGTGTACCAGTTCTTAAGATGCT 149
|||||
552 LeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerLeuLeu 568
|||||
150 CTTGATTACCTGAGAGCAAGATGGAACTGTTTTTCTCTCTCTCTGTT 199
|||||
568 upProThrSerThrSerLeuPheSerTrpTyRAsnSerHisSerSerL 585
|||||
200 GCCAACATCTCTACATCCCTCTCTCTTATTAACAATTCACATGATCTT 249
|||||
585 euSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnGluSer 601
|||||
250 TATCAGTGAATTCCTCAACCAATATTCCTCAGTATTGAACGAGATCA 299
|||||
602 AlavalLeuAlaThrAlaProArgIleaspaspGluIleProProLeu 618
|||||
300 GCTGTACTAGCAACTGCTCCAAAGATAGATGAATCCGCCCTCCACT 349
|||||
618 uproValargThrProGluSerPheIleValValGluGluAlaGlyGluP 635
|||||
350 TCCGTGTCGGACACTGAATCATTTATGTGGTTGAGGAAGCTGGAGAA 399
|||||
635 heSerProAsnValProLysSerLeuSerSerAlaValLysValLysIle 651
|||||
400 TCTCACCAAATGTTCCCAATCCTTATCCTCAGCTGTGAAGGTRAAAT 449
|||||
652 GlyThrSerLeuGluTrpGlyGlyThrSerGluProLysLysPheAspAs 668
|||||
450 GGAACATCAGCTGGAATGGGTGGAACTCTGAACCAAGAAATTTGATGA 499
|||||
668 pSerValIleLeuArgProSerLysSerValLysLeuArgSerPro 683
|||||
500 CTCGTGATATCTAGACCAAGAGTGAATCTCGAAAGTCTCT 545
|||||

```

seq_name: gb_est1:AW504119

seq_documentation_block:

LOCUS AW504119 475 bp mRNA linear EST 02-MAR-2000
DEFINITION UI-HF-BN0-ali-h-10-0-UI.r1 NIH_MGC_50 Homo sapiens cdna clone
IMAGE:3079915 5', mRNA sequence.
ACCESSION AW504119
VERSION AW504119.1 GI:7141786
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 475)
NIH-MGC <http://mgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: M.B. Soares Lab
CDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/brp/image/image.html
Seq primer: MJ3 forward.

FEATURES

Source

```

1. .475
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3079915"
/clone_lib="NIH_MGC_50"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/cell_line="MGC85"

```

```

/lab_host="DH10B (LTI)"
/note="Vector: pT73-Pac; Site_1: NotI; Site_2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(3.5-4.4Kb). Directionally cloned. Cells provided by
Louis M. Staudt, Ph.D. Library preparation by Maria de
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
BASE COUNT 134 a 87 c 115 g 139 t
ORIGIN

```

alignment_scores:

```

Quality: 848.50 Length: 156
Ratio: 5.474 Gaps: 1
Percent Similarity: 99.359 Percent Identity: 99.359

```

alignment_block:

US-09-600-358A-4 x AW504119 ..

Align seg 1/1 to: AW504119 from: 1 to: 475

```

97 LysAlaTyrlleAlaThrGlnGlyProLeuSerThrThrLeuLeuAspPh 113
|||||
10 AAGGCTTATATTGCCACCCAGGGTCTTTATCTACACCCCTCTGGACTT 59
|||||
113 eTrpArgMetIleTrpGluTyRSerValLeuIleIleValMetAlaCysM 130
|||||
60 CTGGAGGATGATTTGGGAATATAGTGTCTTATCATTTGTTATGGCATGCA 109
|||||
130 etGluTyRGlumetGlyLysLysCysGluArgTyRTrpAlaGluPro 146
|||||
110 TGGAGTATGAATGGAAAGAAAAAGTGTAGCGCTACTGGGCTGAGCCA 159
|||||
147 GlyGluMetGlnLeuGluPheGlyProPheSerValSerCysGluAlaGl 163
|||||
160 GGAGAGATGCAGCTGGAATTTGGCCCTTCTCTGTATCTCTGGAAGCTGA 209
|||||
163 uLysArgLysSerAspTyrlleIleArgThrLeuLysValLysPheAsns 180
|||||
210 AAAAGGAATCTGATTATATATATCAGGACTCTAAAGCTTAAGTTCAATA 259
|||||
180 erGluThrArgThrIleTyRGlInPheHisTyRlysAsnTrpProAspHis 196
|||||
260 GTGAACCTCGAATATCTACCAGTTTCATTACAGAAATTGGCCAGACCAT 309
|||||
197 AspValProSerSerIleaspProIleLeuGluLeuIleTrpAspValar 213
|||||
310 GATGTACCTTCTATCTATAGCCCTATTCTTGAGCTCATCTGGGATGTACG 359
|||||
213 qCysTyRGlInGluAspAspSerValProIleCysIleHisCysSerAlaG 230
|||||
360 TTGTTACCAAGAGGATGACAGTGTCCCATATGCAATTCACATGCATGCTG 409
|||||
230 lyCysGlyArgThrGlyValIleCysAlaIleValAspTyRThrTrpMet 246
|||||
410 GCTGTGGAAGGACTGGTGTATTTTGTCTATT...GATTATACATGGATG 456
|||||
247 LeuLeuLysAspGlyIle 252
|||||
457 TTGCTAAAAGATGGGATA 474
|||||

```

seq_name: gb_est1:AUI32040

seq_documentation_block:

LOCUS AUI32040 811 bp mRNA linear EST 24-OCT-2000
DEFINITION AUI32040 NT2RP3 Homo sapiens cdna clone NT2RP303678 5', mRNA
sequence.
ACCESSION AUI32040
VERSION AUI32040.1 GI:10992394
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.


```

451 ThrProPheGluLeuIleGlnArgGluThrLysGluValAspSerLy 467
|||||
365 ACTCCTTTTGATGATACACAGAGAGAAACCAAGGAGGTGGACAGCA 414
|||||
467 sGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheValG 484
|||||
415 GGAAACCTTTTCTTATTGGGATCTCAACACACATGATTCTTTGTTGTAG 464
|||||
484 luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
|||||
465 AGATCGCGGCTCAACAAGTAATGCAATGTTCTTCAACAGAACTGAATAT 514
|||||
501 SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnValLy 517
|||||
515 TCACCTGGCATATGACTCTAACACAGCAATAGCTACTGCTCTACTGTAAAC 564
|||||
517 sHisAspSerSerAlaLeu 524
|||||
565 GCCCATGACTCTAGTGCTCTT 586
|||||

```

seq_name: gb_est1:AW915666

seq_documentation_block:
LOCUS AW915666 554 bp mRNA linear EST 25-MAY-2000
DEFINITION EST346970 Normalized rat embryo, Bento Soares Rattus sp. cDNA clone
RGIC218 5' end, mRNA sequence.

ACCESSION AW915666
VERSION AW915666.1 GI:8081371
KEYWORDS EST.

SOURCE
Rattus sp.

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;

REFERENCE
1 (bases 1 to 554)

AUTHORS
Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,
Kerlavage, A.R. and Adams, M.D.

TITLE
Rat Genome Project: Generation of a Rat EST (RESE) Catalog & Rat

JOURNAL
Unpublished (1998)

COMMENT
Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
This clone is available through the ATCC, contact the ATCC
tel#703-365-2700 for further information
Seq primer: M13 Reverse.

FEATURES
Location/Qualifiers

source
1..554
/organism="Rattus sp."
/db_xref="taxon:10118"
/clone="RGIC218"
/clone_lib="Normalized rat embryo, Bento Soares"
/dev_stage="embryo 8, 12, 18 dpc"
/note="vector: pT7T3pac; site_1: EcoRI; site_2: NotI"

BASE COUNT
ORIGIN

alignment_scores:
Quality: 760.00 Length: 163
Ratio: 4.841 Gaps: 0
Percent Similarity: 96.319 Percent Identity: 86.503

alignment_block:

US-09-600-358A-4 x AW915666 ..

Align seg 1/1 to: AW915666 from: 1 to: 554

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17

```

|||||
66 ATGGACCAAGAGAAATCTCTGCAGCAGCTACTGAAAGAGGCCACAGAAA 115
|||||
17 sLysIleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArgG 34
|||||
116 GAAATTAACAGGGAGGAGTTTGCCCAATGAATTTCTGAAGCTGAAAGGC 165
|||||
34 InSerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50
|||||
166 AATCTACCAAGTACAAAGCAGACAAATATATCTTACAACCGTGGCTCAG 215
|||||
51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
|||||
216 AGGCCCAAGAATATCAAGAAAAATAGATACAGGATATTTTGCCTATGA 265
|||||
67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSerT 84
|||||
266 TCACAGCCTGTAGAGTGTCTCTGTTAACGTCCTGATGAGGACTCCAGTT 315
|||||
84 yrIleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
|||||
316 ACATCAATGCCAGCTTTATTAAAGGCTGCTATGGCCCGCCAGGCTTATATT 365
|||||
101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetil 117
|||||
366 GCCACTCAGGGTCTCTTATCTACAACCTCTCTCGGACTTCTTGAGGATGAT 415
|||||
117 eTrpGluTyrSerValLeuIleValMetAlaCysMetGluTyrGluM 134
|||||
416 CTGGGAATACCGTGTCTGTGGTCATCGTCATGCGCATGTATGGAGTTCGAAA 465
|||||
134 etGlyLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
|||||
466 TGGGAAAGAAAAGTGTGAGCGCTTATTTGGGCTGAGCCAGGAGAAACACAG 515
|||||
151 LeuGluPheGlyProPheSerValSerCysGluAlaGlu 163
|||||
516 CTGCAATTTGGCCCTTTCTATATATCTCTGTGAAGACTGAG 554
|||||

```

seq_name: gb_est2:BG498752

seq_documentation_block:

LOCUS BG498752 745 bp mRNA linear EST 27-MAR-2001
DEFINITION 602544381F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:466653 5',
mRNA sequence.

ACCESSION BG498752

VERSION BG498752.1 GI:13460269

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 745)

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: CLONETECH Laboratories, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCM474 row: a column: 22

High quality sequence stop: 600.

Location/Qualifiers

source

1..745

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:466653"

/clone_lib="NIH_MGC_60"

/tissue_type="adenocarcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);
Site_1: Sfil (ggcgctcgcc); Site_2: Sfil (ggccataggcc
); Double-stranded cDNA was prepared from cell line RNA.
5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGGAGCGGCACATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."

BASE COUNT 245 a 145 c 160 g 195 t
ORIGIN

alignment_scores:
Quality: 738.00 Length: 180
Ratio: 4.761 Gaps: 1
Percent Similarity: 86.111 Percent Identity: 85.000

alignment_block:

US-09-600-358A-4 x BG498752 ..

Align seg 1/1 to: BG498752 from: 1 to: 745

```
1 MetAspGlnArgGluLeuGlnLysPheLeuAspGluAlaGlnSerLy 17
|||||
91 ATGGACCAAGAGAGAAATTCGCAAGAGTTCCTGGATGAGGCCCAAGCAA 140
|||||
17 sLysIleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArg 34
|||||
141 GAAATTAATAAGAGAGGTT.GCCCAATGAATTCGAGCTGAAAGGC 189
|||||
34 InSerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50
|||||
190 AATCTACCAAGTACAGGACAGCAAAACCTATCTACAACTGTGGCTGAG 239
|||||
51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
::: |||||
240 AAGCCCAAGAATAATCAAGAAAAACAGATATAAGGATATTTGCCCTATGA 289
|||||
67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSert 84
|||||
290 TTATAGCCGGTAGAATAATCCCTGATAACCTCTGATGAGGATCCAGCT 339
|||||
84 yTlleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
|||||
340 ACATCAATGCCAACTTCATTAAAGGAGTTATGACCCCAAGGCTTATATT 389
|||||
101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetIl 117
|||||
390 GCCACCCAGGGTCTTATCTACAAACCTCTCTGAGGAGGATGAT 439
|||||
117 eTTPGluTyrSerValIleIleValMetAlaCysMetGluTyrClum 134
|||||
440 TTGGGAATATAGTGTCTTATCATCTGTTATGGCATGCATGGAGTATGAAA 489
|||||
134 etGlyLysLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
|||||
490 TGGGAAG..... 497
|||||
151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysse 167
|||||
498 .....GAAGCTGAAAAAAGGAAATC 517
|||||
167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSer 180
|||||
518 TGATTATATAATCAGGACTCTAAAGTTAAGTTCAATAGT 557
|||||
```

seq_name: gb_est2:BI152499

seq documentation block:

LOCUS BI152499 835 bp mRNA linear EST 05-JUL-2001
DEFINITION 602915640F1 NCL_CGAP_Lu29 Mus musculus cDNA clone IMAGE:5066337 5',
mRNA sequence.
ACCESSION BI152499
VERSION BI152499.1 GI:14612500
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 835)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11178 row: o column: 10

High quality sequence stop: 762..

FEATURES

source

1..835 Location/Qualifiers

/organism="Mus musculus"

/strain="C57BL/6J (f1)"

/db_xref="taxon:10090"

/clone="IMAGE:5066337"

/clone_lib="NCI_CGAP_Lu29"

/tissue_type="spontaneous tumor, metastatic to mammary."

Stem cell origin.

/lab_host="DH10B"

/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: SalI;

Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Library constructed by Life Technologies. Investigator

providing samples: Gilbert Smith, NIH"

BASE COUNT 261 a 151 c 210 g 213 t

ORIGIN

alignment_scores:

Quality: 718.50 Length: 304

Ratio: 3.421 Gaps: 7

Percent Similarity: 69.079 Percent Identity: 49.671

alignment_block:

US-09-600-358A-4 x BI152499 ..

Align seg 1/1 to: BI152499 from: 1 to: 835

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17

|||||

7 ATGGACCAAGTGGAGATCCCTGAGGAGGTTTCATGGAGAGGTCAGGCCAT 56

|||||

17 sLysIleThr.....LysGluGluPheAlaAsnGluPheLeuL 30

|||||

57 GAAGAGTCCGACCACCAATGGGAGGAGCAACTGCCCGGACATCATGT 105

|||||

30 ysLeuLysArgGlnSerThrLysTyrLysAlaAspLysThrTyrProThr 46

|||||

106 T..... 107

47 ThrValAlaGluAsnAlaLysAsnIleLysLysAsnArgTyrLysAspIl 63

|||||

107 107

63 eLeuProTyrAspTyrSerArgValGluLeuLeuSerLeuIleThrSerAspG 80

|||||

.....

108GATCACAGCCGAGTTAGTGTGCATTGAAGATTTCATCCC 147

80 luAspSerSertYrileAasnAlaaSnPheIleLysGLyValtyrdlyPro 96
::|||||::|||::|||::|||::|||::|||::|||::|||::|||
148 AAGATTcagAtTAATACTCAATCGCAAATTTAAAGAAGGGTGCTATGGGCCA 197

97 LysalaTyrlleaIAthrGIinglYproLeuserrThrLeuLeuaSpH 113
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
198 AAAGCATATGTGGCAACCACAAGGGCCCTTGGCAATACAGTCATAGACTT 247

113 eTrpArgMetIetrpGlutyrSerValLeuilleleIValMetalacYSm 130
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
248 CTGGAGGATGATATGGGAGATAATATGTGTGCTATATCGTAGCGCTGTC 297

130 etGluTYrGluMeCtLYlsLysCysGLuarGYrrTpAlaGluPro 146
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
298 GAGAAATTTGAGATGGGAAG..... 317

147 GlyglumetGINLeuGluPhesGlyProPheSerValSerCysGluAlagl 163
||| || |
318GAAAATGA 325

163 uLySArgLYSSerASptYrIllelleArqThrLeuLysVallysPheAnsN 180
:||: ::::::|||::|||::|||::|||::|||::|||::|||::|||
326 ACAAGCAAGAACCCAGCTACTTTCATCCGAACACTTTTTACTTGAATTTCAA 375

180 erGluTHrARGThRIletYrcInPheHisTYrLYSasNrtproASPHeIS 196
:::||||::|||::|||::|||::|||::|||::|||::|||::|||
376 ATGAATCCCCGTCGCTCTATCAGTTTCATTACGTGAACCTGGCCAGCACCAT 425

197 AspValPröserSeRilEaspProilleuLeuLUeulleITrpApSValar 213
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
426 GATGTTCTTCGTCAATTGATTCATTCTGGACATGATAAGCTTAAATGAG 475

213 gCYSTYrcINdlnuSPspSrvalPROilEcysIIehis..CyssERAla 229
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
476 GAATACCAAGAACATGAAGATGTCCCTATTGTTGATTTCATTTCAGTAGCA 525

230 GLycSYSLYGrtHRgLYVaILecYSALatlieVALaspTYrTHRTPme 246
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
526 GGCTGTGGCAACAGGTGCTATTTTGCCATA...GATTACACGTGGNA 572

246 tLeuLeuLYSaSPglyllleIEProGLunASNpheSerValPheSerLeui 263
||||| || |
573 CTTACTGAAGACGGAAAAATTCCAGAGGAATTTAACTATTATTAATTTAA 622

263 leARGLUmetARGThRClnARGProSERLeuValGlnThRClnGluGln 279
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
623 TACAAGAAATGAGAACACAGAGGCCTCGGCAGTACAAACAAGAGGAGCAG 672

280 TyRGluLeuValTYr.AsnAlaValLeuGluLeupHelLYsarGGINMet 295
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
673 TATGAAC TTGGTTCATAGAAGCTATTGCTCAACTGTTTGCAAAAACAGGTA 722

296 AspVal 297
::: || |
723 CAAC TG 728

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2002, 08:00:48 ; Search time 474.64 Seconds
(without alignments)
8522.355 Million cell updates/sec

Title: US-09-600-358a-3
Perfect score: 2356

Sequence: 1 tccctcaacctacttataga.....cctagtcggagcacatccgg 2356

Scoring table: IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT.*
2: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT.*
3: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT.*
4: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT.*
5: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT.*
6: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT.*
7: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT.*
8: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT.*
9: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT.*
10: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT.*
11: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.*
12: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT.*
13: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT.*
14: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT.*
15: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT.*
16: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT.*
17: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT.*
18: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT.*
19: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT.*
20: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT.*
21: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.*
22: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT.*
23: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT.*
24: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2356	100.0	2356	20	Human Cytoplasmic
2	2097	89.0	3058	20	Human Cytoplasmic
3	2074.6	88.1	3580	20	Human PTP04 encodi
4	1729	73.4	2452	22	Human tyrosine pho
5	621.4	26.4	1286	21	Human breast and o
6	368	15.6	2983	24	Mouse ischaemic co
7	216.4	9.2	1529	18	cDNA encoding a no
8	214.2	9.1	2226	19	Rat protein tyrosi
9	202.2	8.6	2810	19	Human brain derive

10	160	6.8	205	21	AA012676	Human secreted pro
11	138.8	5.9	4798	22	AA158418	Human polynucleoti
12	130.6	5.5	2309	15	AA073786	Partial PCR fragme
13	129	5.5	4965	22	AA098348	Human EST-derived
14	129	5.5	4968	22	AA160204	Human polynucleoti
15	120	5.1	466	18	AA03128	Partial cDNA encod
16	116.8	5.0	5117	16	AA06027	huDEP-1 cDNA. Hom
17	115.4	4.9	3969	19	AAV05762	Chicken protein ty
18	115.4	4.9	3969	19	AAV81897	Chicken transmembr
19	115.2	4.9	2674	22	AA099616	Human protein enco
20	115.2	4.9	5069	22	AAK52292	Human polynucleoti
21	115.2	4.9	5164	22	AAK53276	Human polynucleoti
22	115.2	4.9	5465	22	AAK50778	Human cancer agent
23	112.8	4.8	4555	16	AA094311	Tyrosine phosphata
24	106.2	4.5	247	18	AA069301	Murine metastatic
25	104.8	4.4	4752	19	AAV04868	Nucleotide sequenc
26	104.8	4.4	4799	19	AAV04887	Nucleotide sequenc
27	104.8	4.4	4815	19	AAV04886	Nucleotide sequenc
28	104.8	4.4	5113	17	AA048479	Human tyrosine dep
29	104	4.4	2760	24	AB199784	Mouse ischaemic co
30	100.6	4.3	5455	19	AAV33577	Osteosteticular pr
31	98.4	4.2	4651	15	AA072913	Human receptor-tyr
32	98.4	4.2	4823	23	AAK52612	DNA encoding novel
33	98.4	4.2	5775	20	AAZ08539	Human protein phos
34	92.4	3.9	6075	21	AA088866	Human protein tyro
35	92.4	3.9	6342	22	AAK51954	Human polynucleoti
36	92.4	3.9	6346	22	AAK52938	Human polynucleoti
37	92.4	3.9	6632	22	AA026742	Human breast cance
38	90.4	3.8	3623	20	AAV82821	Rabbit osteoclast-
39	90.4	3.8	4573	22	AA098379	Human EST-derived
40	90.4	3.8	5412	24	AB199920	Rat mucocardial ce
41	90.4	3.8	5690	16	AA086902	Rat receptor type-
42	90	3.8	6000	16	AA086478	Human PTP-OB. Hom
43	90	3.8	6000	18	AA085389	Human protein tyro
44	90	3.8	6000	20	AAK06095	Human protein tyro
45	90	3.8	6000	22	AA02144	Human protein tyro

ALIGNMENTS

RESULT 1

AA090696

ID AA090696 standard; cDNA; 2356 BP.

AC AA090696;

XX 01-OCT-1999 (first entry)

XX Human Cytoplasmic phosphatase Lyp2, cDNA.

DE Lymphoid Protein Tyrosine Phosphatase gene; Lyp gene; immunosuppressant;

KW intracellular tyrosine phosphatase; Ptpase; lymphocyte; fetal liver;
KW resting lymphoid cell; protein tyrosine kinase; PTKs; transgenic animal;
KW lymphoma; T cell antigen receptor signalling; intronic sequence;
KW cytokine receptor signalling; protein tyrosine phosphatase; Ptpase; ds.
XX Homo sapiens.
OS

XX Key Location/Qualifiers

FT CDS 42..2120

FT /*tag= a

FT /product= "Cytoplasmic phosphatase, Lyp2"

FT /note= "Expressed in fetal liver tissue and resting lymphoid cells"

FT WO9936548-A1.

PN 22-JUL-1999.

XX 18-JAN-1999; 99WO-CA00038.

XX 16-JAN-1998; 98CA-2220853.

PR

QY	1621	attcttacatacctttagtgaaaaatcccttatttttcatcatggtccccaagtgggtacca	1680
Db	1621	attcttacatacctttagtgaaaaatcccttatttttcatcatggtccccaagtgggtacca	1680
QY	1681	gttctaagaatgctcttgatttaccctgagaagaagaatggaactgttttttccttctctc	1740
Db	1681	gttctaagaatgctcttgatttaccctgagaagaagaatggaactgttttttccttctctc	1740
QY	1741	tgttgccaacatccctacatccctctctctcttattacaattcaacatgcttttttatacc	1800
Db	1741	tgttgccaacatccctacatccctctctctcttattacaattcaacatgcttttttatacc	1800
QY	1801	tgaattctccaaccaatatttctcactatgaaaccaggagtcagctgttactagaacctg	1860
Db	1801	tgaattctccaaccaatatttctcactatgaaaccaggagtcagctgttactagaacctg	1860
QY	1861	ctccaagatagatgaataatccccctccacttctctgtacggacacactggaatacttta	1920
Db	1861	ctccaagatagatgaataatccccctccacttctctgtacggacacactggaatacttta	1920
QY	1921	ttgtggtgaggaagctggagaattctcaccaaatgttccaaatccttatctcctcagctg	1980
Db	1921	ttgtggtgaggaagctggagaattctcaccaaatgttccaaatccttatctcctcagctg	1980
QY	1981	tgaaggtaaaaatttggaacatcactgggaatggggtggaacatctgaaccacaaagaatttg	2040
Db	1981	tgaaggtaaaaatttggaacatcactgggaatggggtggaacatctgaaccacaaagaatttg	2040
QY	2041	atgactctgtgatacttagaccacgaagaagtgtaaaactccgaagtctcaaatccggta	2100
Db	2041	atgactctgtgatacttagaccacgaagaagtgtaaaactccgaagtctcaaatccggta	2100
QY	2101	aaaattctcttgcttttagatgacatttagccctaaagattggaagaatggttcgttaag	2160
Db	2101	aaaattctcttgcttttagatgacatttagccctaaagattggaagaatggttcgttaag	2160
QY	2161	tttagagtaattcacttcagggaagttaactgtgttcccataatgacttccagttattcatg	2220
Db	2161	tttagagtaattcacttcagggaagttaactgtgttcccataatgacttccagttattcatg	2220
QY	2221	atttattcttgctttcccagactagaaattttgtaaagagtcatagggggaagctagggc	2280
Db	2221	atttattcttgctttcccagactagaaattttgtaaagagtcatagggggaagctagggc	2280
QY	2281	taaccagaaaaataaaataatggataaaaaatcggaactactgttttcccccta	2340
Db	2281	taaccagaaaaataaaataatggataaaaaatcggaactactgttttcccccta	2340
QY	2341	gtcggagcacatcccg 2356	
Db	2341	gtcggagcacatcccg 2356	
RESULT 2			
AAX90695			
ID	AAX90695 standard; cDNA; 3058 BP.		
XX	AAX90695;		
AC			
XX			
DT	01-OCT-1999 (first entry)		
XX			
DE	Human Cytoplasmic phosphatase Lyp 1, CDNA.		
XX			
KW	Lymphoid Protein Tyrosine Phosphatase gene; Lyp gene; lymphoid cell;		
KW	intracellular tyrosine phosphatase; ptpase; lymphocyte; thymocyte;		
KW	T cell; B cell; protein tyrosine kinase; ptk; immunosuppressant;		
KW	lymphoma; transgenic animal; T cell antigen receptor signalling;		
KW	cytokine receptor signalling; protein tyrosine phosphatase; ptpase; ds		
XX	Homo sapiens.		
OS			
XX			
FH	Key	Location/Qualifiers	

PR 20-MAY-1997; 97US-0047222.
PR 11-JUN-1997; 97US-0049477.
PR 11-JUN-1997; 97US-0049756.
PR 18-JUN-1997; 97US-0049914.
XX
FA (SUGE-) SUGEN INC.
XX
XX App H, Clary D, Courtneidge SA, Hui TH, Jallal B;
PI Markby D, Onrust S, Peles E, Plowman GD;
XX
XX WPI: 1999-009434/01.
DR P-PSDB: AAW89247.
DR
XX New nucleic acid encoding specific protein tyrosine phosphatases -
PT useful for identifying specific modulators for treatment and
PT prevention of cancer and neurodegenerative disease
XX
XX Claim 2; Page 145-146; 193pp; English.
XX
XX The present invention describes isolated, enriched or purified nucleic
CC acids encoding PTP04, SAD, PTP05, PTP10, ALP and ALK-7 proteins. The
CC present sequence encodes human PTP04. The above proteins, other than
CC ALK-7, are protein tyrosine phosphatases (PTPs) and are used to identify
CC substances that modulate their activity (i.e. agonists and antagonists,
CC including NBP) in vivo or in vitro. These substances are used to treat
CC or prevent diseases associated with abnormal signal transduction
CC pathways that involve the proteins, particularly cancer (e.g. leukemia
CC and lymphoma), while modulators of ALK-7 (which is a type I receptor
CC serine/threonine kinase) are used to promote neuronal survival,
CC particularly for treating Alzheimer's, Parkinson's or Huntington's
CC diseases. Nucleic acid fragments of the polynucleotides encoding the
CC proteins can be used as probes to identify and clone related sequences;
CC to detect protein-encoded RNA; to generate transgenic animals and in
CC gene therapy (optionally after mutation). Ab are used to determine the
CC proteins.
XX
XX Sequence 3580 BP; 1213 A; 666 C; 618 G; 1083 T; 0 other;

Query Match 88.18; Score 2074.6; DB 20; Length 3580;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 2090; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

Qy 1 tccctcaacctactatagactattttctgctcgcagcatgagcacaagagaaattc 60
Db 12 tccctcaacctactatagactattttctgctcgcagcatgagcacaagagaaattc 71
Qy 61 tgcagaagttcctgagtgagcccaagcaagaaattactaaagagaggttgccaatg 120
Db 72 tgcagaagttcctgagtgagcccaagcaagaaattactaaagagaggttgccaatg 131
Qy 121 aatttctgaagctgaaaggaattctaccagtaacagagcaagcaaaacctatccaa 180
Db 132 aatttctgaagctgaaaggaattctaccagtaacagagcaagcaaaacctatccaa 191
Qy 181 ctgtggtcgaagatgccaagaaatcaagaaacagagataaaggtatttgccctatg 240
Db 192 ctgtggtcgaagatgccaagaaatcaagaaacagagataaaggtatttgccctatg 251
Qy 241 attatgcccgggtagaactatccctgataacctctgtagaggtattccagctacatcaatg 300
Db 252 attatgcccgggtagaactatccctgataacctctgtagaggtattccagctacatcaatg 311
Qy 301 ccaacttcaataagggaggtatttgacccaagggctattatgccaacccaggtcccttat 360
Db 312 ccaacttcaataagggaggtatttgacccaagggctattatgccaacccaggtcccttat 371
Qy 361 ctacaacctctgactctcggagatgatttgggaataagtagtcccttatcatgttta 420
Db 372 ctacaacctctgactctcggagatgatttgggaataagtagtcccttatcatgttta 431
Qy 421 tggcatgcatgagtatgaaatggaaagaaagagtgagcgctactgggctgagccag 480
Db

Db 432 tggcatgcatgagtatgaaatggaaagaaagtgagcgctactgggctgagccag 491
Qy 481 gagagatcagctgaaatttgcccttctctgtatctctgtaagctgaaagcaaaagaaat 540
Db 492 gagagatcagctgaaatttgcccttctctgtatctctgtaagctgaaagcaaaagaaat 551
Qy 541 ctgattataatcagagactctaaagttaagttaagttaagttaagttaagttaagttaagt 600
Db 552 ctgattataatcagagactctaaagttaagttaagttaagttaagttaagttaagttaagt 611
Qy 601 agttctatacaagaatttggccagaccatgatgtaccctctcatctatagaccctattcttg 660
Db 612 agttctatacaagaatttggccagaccatgatgtaccctctcatctatagaccctattcttg 671
Qy 661 agtcatctggatgacgttgttaccagagagatgacagtggttcccatatgcatcact 720
Db 672 agtcatctggatgacgttgttaccagagagatgacagtggttcccatatgcatcact 731
Qy 721 gcagtctggcttgggaaggactggtgtatttggctatttggatttatcatcatggtatgt 780
Db 732 gcagtctggcttgggaaggactggtgtatttggctatttggctatttatcatggtatgt 788
Qy 781 tgcataaagatgggaataattcctcgagaactctgaagcttctgagcttctgagcttctgag 840
Db 789 tgcataaagatgggaataattcctcgagaactctgaagcttctgagcttctgagcttctgag 848
Qy 841 ggacacagagcccttcattagttcaaacgcaggaacaatatgaactggtctcacaatgctg 900
Db 849 ggacacagagcccttcattagttcaaacgcaggaacaatatgaactggtctcacaatgctg 908
Qy 901 tattagaactatttaagagacagatggatgttatcagagataaacattctcggaaacagaga 960
Db 909 tattagaactatttaagagacagatggatgttatcagagataaacattctcggaaacagaga 968
Qy 961 gtcaagcaaaacattgtattcctcgagaataatcacactctccaaagcagactcttattctc 1020
Db 969 gtcaagcaaaacattgtattcctcgagaataatcacactctccaaagcagactcttattctc 1028
Qy 1021 ctaattttaccaaaagtagtaccacaaagcagcaaaaatgatgaaccaacaaagagcaaaa 1080
Db 1029 ctaattttaccaaaagtagtaccacaaagcagcaaaaatgatgaaccaacaaagagcaaaa 1088
Qy 1081 tggaaatcaaaagatcttctctcttctgacttttagactctctgaaataagtgcaaaagaag 1140
Db 1089 tggaaatcaaaagatcttctctcttctgacttttagactctctgaaataagtgcaaaagaag 1148
Qy 1141 agctagtttgcacctgctaaatcaagcactcttctgactttctgactttctgagctaaattaca 1200
Db 1149 agctagtttgcacctgctaaatcaagcactcttctgactttctgactttctgagctaaattaca 1208
Qy 1201 gttttgacaaaatgctgacacaaacctgaaatggcagacaaagcatttccaatagtttg 1260
Db 1209 gttttgacaaaatgctgacacaaacctgaaatggcagacaaagcatttccaatagtttg 1268
Qy 1261 gggagcctcttcagaagcatcaaaagtgtggatttgggtctctcttcttcttggaggatgtt 1320
Db 1269 gggagcctcttcagaagcatcaaaagtgtggatttgggtctctcttcttcttggaggatgtt 1328
Qy 1321 ctaattctaaacctgtaaatgcagcaggaagataatttaattcaaaagtgccaataaacac 1380
Db 1329 ctaattctaaacctgtaaatgcagcaggaagataatttaattcaaaagtgccaataaacac 1388
Qy 1381 ggacaaaatcaacctcttcttgaattgatacagcagagaaacaaagaggtggagacaca 1440
Db 1389 ggacaaaatcaacctcttcttgaattgatacagcagagaaacaaagaggtggagacaca 1448
Qy 1441 agggaaactcttcttcttgaattgatacagcagagaaacaaagaggtggagacaca 1500
Db 1449 agggaaactcttcttcttgaattgatacagcagagaaacaaagaggtggagacaca 1508
Qy 1501 ctcaaaaagtaagtcatgtttcttcagcagaactgaaattattcactgccaatgactcta 1560
Db 1509 ctcaaaaagtaagtcatgtttcttcagcagaactgaaattattcactgccaatgactcta 1568

QY 1561 aacacaaataacgaatgctcttaataatgtaaaagcacacatgactctagtctcttggtgtat 1620
|||||
Db 1569 aacacaaataacgaatgctcttaataatgtaaaagcacacatgactctagtctcttggtgtat 1628
|||||
QY 1621 attcttacatcacttttagtggaaaaatccctatttttttccatcgtgcctccaagtgggtacca 1680
|||||
Db 1629 attcttacatcacttttagtggaaaaatccctatttttttccatcgtgcctccaagtgggtacca 1688
|||||
QY 1681 gttctaaagtgtctcttgatttaccctgagaagcaaatggaactgttttttctctcttc 1740
|||||
Db 1689 gttctaaagtgtctcttgatttaccctgagaagcaaatggaactgttttttctctcttc 1748
|||||
QY 1741 tgtgtccaaacatctctacatccctctctcttctattacaaattcacatagttctttttatcac 1800
|||||
Db 1749 tgtgtccaaacatctctacatccctctctcttctctattacaaattcacatagttctttttatcac 1808
|||||
QY 1801 tgaattctccaaacaaatattctctcactattgaaccagagtcagctgtactagcaactg 1860
|||||
Db 1809 tgaattctccaaacaaatattctctcactattgaaccagagtcagctgtactagcaactg 1868
|||||
QY 1861 ctccaaggtatagatgataaattccccctccacttctctgtacggacacactgaaatcttta 1920
|||||
Db 1869 ctccaaggtatagatgataaattccccctccacttctctgtacggacacactgaaatcttta 1928
|||||
QY 1921 ttgtggttgagggaagctggagaattctccacaaatgttcccaaatctcttctctcagctg 1980
|||||
Db 1929 ttgtggttgagggaagctggagaattctccacaaatgttcccaaatctcttctcagctg 1988
|||||
QY 1981 tgaagtgaaaaattgaacatacactggaattggggtggaacatctcgaaccaaagaaatttg 2040
|||||
Db 1989 tgaagtgaaaaattgaacatacactggaattggggtggaacatctcgaaccaaagaaatttg 2048
|||||
QY 2041 atgactctgtactactaagcaagcaagagtgtaaaactccgaactcgaactcgaactcga 2097
|||||
Db 2049 atgactctgtactactaagcaagcaagagtgtaaaactccgaactcgaactcgaactcga 2105
|||||

RESULT 4
AAH79342
ID AAH79342 standard; DNA; 2452 BP.
XX
AC AAH79342;
XX
XX
DT 04-DEC-2001 (first entry)
XX
DE Human tyrosine phosphatase coding sequence.
XX
KW Human; tyrosine phosphatase; hPTP; ds.
XX
OS Homo sapiens.
XX
PN CN1302899-A.
XX
PD 11-JUL-2001.
XX
PE 29-OCT-1999; 99CN-0119935.
XX
PR 29-OCT-1999; 99CN-0119935.
XX
PA (NANF-) NANFANG RES CENT STATE HUMAN GENE GROUP.
XX
PI Ren S, Wu T, Qian B;
XX
XX WPI; 2001-550591/62.
DR P-PSDB; AAG78623.
PT Human protein tyrosine phosphatase and its coding sequence -
XX
XX Claim 1; Page 19-21(Disclosure); 29pp; Chinese.
XX
XX The present invention provides the protein and coding sequences of human
CC tyrosine phosphatase hPTP. The protein is expressed in human normal
CC

CC suprarenal tissue. The present sequence is the coding sequence of the
XX invention.
SQ Sequence 2452 BP; 791 A; 516 C; 468 G; 677 T; 0 other;

Query Match 73.4%; Score 1729; DB 22; Length 2452;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1946; Conservative 0; Mismatches 100; Indels 27; Gaps 13;

QY 33 ctctgcagcatgaccacaaagaaaaattctgcagaagtctctctggtatgagcccaaaagcaag 92
Db 1 ctctgcagcatgaccacaaagaaaaattctgcagaagtctctctggtatgagcccaaaagcaag 60
QY 93 aaaaattactaaagagaggtttgccaatgaattcttgaagctgaaaaggaatctaccag 152
Db 61 aaaaattactaaagagaggtttgccaatgaattcttgaagctgaaaaggaatctaccag 120
QY 153 tacaaggcgacaaaaacctctctacaactgtgctgagaatgccagaataatcaagaaa 212
Db 121 tacaaggcgacaaaaacctctctacaactgtgctgagaatgccagaataatcaagaaa 180
QY 213 aacagataaagatattttgcccctatgattatagccgggtgagaactatccctgataacc 272
Db 181 aacagataaagatattttgcccctatgattatagccgggtgagaactatccctgataacc 240
QY 273 tctgatgagattcccagctcacatcaatgccaaacttcattaaaggagtttatgacccaag 332
Db 241 tctgatgagattcccagctcacatcaatgccaaacttcattaaaggagtttatgacccaag 300
QY 333 gcttatattgccaccacaggggtctcttatctacaacctctctgacttctgagagatgatt 392
Db 301 gcttatattgccaccacaggggtctcttatctacaacctctctgacttctgagagatgatt 360
QY 393 tgggaataatagtgctcattatgattgcatgcatgagatgataaaatgggaaagaaa 452
Db 361 tgggaataatagtgctcattatgattgcatgcatgagatgataaaatgggaaagaaa 420
QY 453 aagtgtagcgtactggctgagccagagagatgcagctggaatttgccctctctct 512
Db 421 aagtgtagcgtactggctgagccagagagatgcagctggaatttgccctctctct 480
QY 513 gtatcctgtgaagctgaaaaaagaaaatctgtattataatcaggagactctaaagttaag 572
Db 481 gtatcctgtgaagctgaaaaaagaaaatctgtattataatcaggagactctaaagttaag 540
QY 573 ttcaatagtgaaactcgaactatctaccagtttcattacaagaattggccagaccatgat 632
Db 541 ttcaatagtgaaactcgaactatctaccagtttcattacaagaattggccagaccatgat 600
QY 633 gtacctctcatatagaccctattcttgagctcatctgggatgtacgtgtgttaccagag 692
Db 601 gtacctctcatatagaccctattcttgagctcatctgggatgtacgtgtgttaccagag 660
QY 693 gatgacagtggtcccatatgcattcactgctgctggctggtggaaggagactggtgtatt 752
Db 661 gatgacagtggtcccatatgcattcactgctgctggctggtggaaggagactggtgtatt 720
QY 753 tgtgctattgttattatatacatgattgtctaaaaagatgggataattctctgagaacttc 812
Db 721 tgtgcta---ttgattatacatgattgttctaaaaagatgggataattctctgagaacttc 777
QY 813 agtgttttcagtttgatccgggaaatcggcacacagagcccttcattagttcacaacagag 872
Db 778 agtgttttcagtttgatccgggaaatcggcacacagagcccttcattagttcacaacagag 837
QY 873 gaacaatatgaactggtctacaatgctgtatttagaactatttaagagacagatggtatggt 932
Db 838 gaacaatatgaactggtctacaatgctgtatttagaactatttaagagacagatggtatggt 897
QY 933 atcagagataaaacattcttgaaacagagatgcaagcaaacattgtattctctgagaaaaat 992
Db 898 atcagagataaaacattcttgaaacagagatgcaagcaaacattgtattctctgagaaaaat 957

```
QY 993 cacactctcaagcagactcttattctcttaatttaccataaagtaccacaaagcagca 1052
Db 958 cacactctcaagcagactcttattctcttaatttaccataaagtaccacaaagcagca 1017
QY 1053 aaaaatgatgaacacaaagcagcaaaaatggaatacaaaatcttctcttctgacttt 1112
Db 1018 aaaaatgatgaacacaaagcagcaaaaatggaatacaaaatcttctcttctgacttt 1077
QY 1113 aggaactctgaaataagtgcaaaaagagagctagttttgcacccctgctaaatcaagcact 1172
Db 1078 aggaactctgaaataagtgcaaaaagagagctagttttggcccccgg---aaaaccaggct 1134
QY 1173 tcttttgacttctgagctgaaataatgagtaatttgacaaaatgctgacacaaaccagaaa 1232
Db 1135 tcttttgacttctgagctgaaataatgagtaatttgagtttg---caaaaatgctgcccacacagaaa 1192
QY 1233 tggcagacaaagcatttccaaatagttgggagcctcttcagaaagcatcaaaagtgtggat 1292
Db 1193 tggcag-caaaagccttccaaatagt--gggaagccttccagaaagcatcaaaag-ttgaat 1248
QY 1293 tgggctctctttgtttgagggatgtcttaattctaaacctgtaaatgcagcagagaaga 1352
Db 1249 tgggctctctttgtttt---aggaatgtcttaattctaaacctgt--atgcagcagagaaga 1303
QY 1353 tatttaattcaaaagtgccaataacacgcagacaaatacactccttttgaattgatcacag 1412
Db 1304 tattt--aatccaagtgccaat-accgagacaaatacact-cctttgaattgataccag 1359
QY 1413 cagagagaacaaagaggtggacagcaaggaatactttcttatttggaaatcacaacca 1472
Db 1360 ca---agagaacaaaggtgacagcaagaaaa---ccttctatttgaatcacaacca 1413
QY 1473 catgattcttgtttgtagagatgcaggctcaaaaagtaagtaatgtttttcagcagaa 1532
Db 1414 catgattcttgtttgtagagatgcaggctcaaaaagtaagtaatgtttttcagcagaa 1473
QY 1533 ctgaattattcactgcataatgacttaacacccaataacgtaatgcctctaagttaaaag 1592
Db 1474 ctgaattattcactgcataatgacttaacacccaataacgtaatgcctctaagttaaaag 1533
QY 1593 caccatgactctagtgctcttctgtgtatattcttatacatccttttagtggaaaatccttat 1652
Db 1534 caccatgactctagtgctcttctgtgtatattcttatacatccttttagtggaaaatccttat 1593
QY 1653 ttttcatatgctcccaagtggtaccaggtcttaagatgtctctgtatttaccctgagaag 1712
Db 1594 ttttcatatgctcccaagtggtaccaggtcttaagatgtctctgtatttaccctgagaag 1653
QY 1713 caagatggaactgttttctctctctctgtgtcccaacatcctctacatccctctctct 1772
Db 1654 caagatggaactgttttctctctctctgtgtcccaacatcctctacatccctctctct 1713
QY 1773 tattacaattcacatagttcttttatacactgaattctcccaacaaatatttctcactattg 1832
Db 1714 tattacaattcacatagttcttttatacactgaattctcccaacaaatatttctcactattg 1773
QY 1833 aaccaggagtcagctgtactagcaactgtctcccaagatagatgataaaatccccctccca 1892
Db 1774 gaacaggagtcagctgtactagcaactgtctcccaagatagatgataaaatccccctccca 1833
QY 1893 ctctcctgtacgacactgaactattatttattgtgttgaggaaagctgagaattctcacca 1952
Db 1834 tctccttaaggggacactgaactattatttattgtgttgaggaaagctgagaattctcacca 1893
QY 1953 aatgttccccaaatccttactcctcagctgtgaaaggttaaaaattggaacatacactgggaatg 2012
Db 1894 aatgttccccaaatccttactcctcagctgtgaaaggttaaaaattggaacatacactgggaatg 1953
QY 2013 ggttggaactctgaaacaaagaaattttagtgactctgtgatacttagaccaaagcaagat 2072
Db 1954 ggttggaactctgaaacaaagaaattttagtgactctgtgatacttagaccaaagcaagat 2013
```

```
QY 2073 gtaaaactcgaagctctctaaatcagtgtaaaaaat 2105
Db 2014 gtaaaactcgaagctctctaaatcagtgtaaaaaat 2046

RESULT 5
ID AAF21801
XX AAF21801 standard; DNA; 1286 BP.
AC AAF21801;
XX
DT 27-MAR-2001 (first entry)
XX
DE Human breast and ovarian cancer associated antigen gene SEQ ID 188.
XX
KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW neotropic; neurprotection; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antinflammatory; antitumor; antiparasitic; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease; ds.
XX
OS Homo sapiens.
XX
XX WO200055173-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US05881.
XX
XX 12-MAR-1999; 99US-0124270.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI: 2000-611515/58.
XX
XX P-PSDB; AAB58898.
XX
XX New human breast and ovarian cancer associated gene sequences and the
XX polypeptides encoded by these genes, useful in the prevention,
XX treatment and diagnosis of cancer, immune disorders, cardiovascular
XX disorders and neurological diseases
XX
XX Claim 1; Page 622; 1299pp; English.
XX
XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
XX proteins AAB58711 - AAB59128. The DNA and protein sequences are
XX associated with breast and ovarian cancer. Included in the invention are
XX sequences AAF22032 - AAF22040 and AAB59129 which are used in the
XX isolation and characterisation of the DNA and protein sequences of the
XX invention. The breast and ovarian cancer associated DNA, protein, agonist
XX or antagonist sequences exhibit cytostatic; immunosuppressive;
XX neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
XX antidiabetic; antinflammatory; antitumor; antiparasitic and cardiac activity. The
XX antibacterial; antifungal; antiparasitic and anticonvulsant activity of cancer,
XX polynucleotide and protein sequences are used in the diagnosis of cancer,
XX particularly breast and ovarian cancer. The nucleic acid sequences,
XX proteins, agonists and antagonists may also be used in the diagnosis,
XX prevention and treatment of immune disorders e.g. Addison's disease,
XX allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
XX diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
XX arthritis and ulcerative colitis; cardiovascular disorders such as
XX myocardial ischaemias; wound healing; neurological diseases such as
XX cerebral anoxia and epilepsy; and infectious diseases.
XX
XX Sequence 1286 BP; 410 A; 273 C; 209 G; 389 T; 5 other;
XX

Query Match 26.4%; Score 621.4; DB 21; Length 1286;
Best Local Similarity 98.7%; Pred. No. 4.2e-154;
```


Matches 625; Conservative 1; Mismatches 7; Indels 0; Gaps 0;			
Qy	1473	catgattctgtttgtgagatgcaggctcaaaaagtaatgcattgttcttcagcagaa	1532
Db	2	catgattctgtttgtgagatgcaggctcaaaaagtaatgcattgttcttcagcagaa	61
Qy	1533	ctgaattattcactgccatgactctaaacacccaatacgttaatgcctctaataag	1592
Db	62	ctgaattattcactgccatgactctaaacacccaatacgttaatgcctctaataag	121
Qy	1593	caccatgactctagtcgttgggtatattcttaccatccttttagtggaaaaactctat	1652
Db	122	caccatgactctagtcgttgggtatattcttaccatccttttagtggaaaaactctat	181
Qy	1653	titttcatacgtccctcccaagtgaccagttctaagatgtccttgattaccctgagaag	1712
Db	182	titttcatacgtccctcccaagtgaccagttctaagatgtccttgattaccctgagaag	241
Qy	1713	caagatggaaactgttttctctctctctgttgcccaacatcctctacatccctctctct	1772
Db	242	caagatggaaactgttttctctctctctgttgcccaacatcctctacatccctctctct	301
Qy	1773	tattacaattcacatgattctttatcactgaattctccaacaaatattctcactattg	1832
Db	302	tattacaattcacatgattctttatcactgaattctccaacaaatattctcactattg	361
Qy	1833	aaccaggagtcagctgactagcaactgctccaagatagatgataatccccctcca	1892
Db	362	aaccaggagtcagctgactagcaactgctccaagatagatgataatccccctcca	421
Qy	1893	cttctctgacggacacacctgaatcatttattgttggtggaaagctggagaattctcacca	1952
Db	422	cttctctgacggacacacctgaatcatttattgttggtggaaagctggagaattctcacca	481
Qy	1953	aatgttcccaaatccttactcctcagctgtaagggttaaaaattggaaacatcactggaatgg	2012
Db	482	aatgttcccaaatccttactcctcagctgtaagggttaaaaattggaaacatcactggaatgg	541
Qy	2013	ggtggaacatctgaacccaagaatttgatgactctgtgatacttagaccagcagaagt	2072
Db	542	ggtggaacatctgaacccaagaatttgatgactctgtgatacttagaccagcagaagt	601
Qy	2073	gtaaaactcgaagtctcaataatcaggtaaaaat	2105
Db	602	gtaaaactcgaagtctcaataatcagaactacat	634
RESULT 6			
ABI99910			
ID	ABI99910	standard; cDNA; 2983 BP.	
AC	ABI99910;		
XX			
DT	07-MAR-2002	(first entry)	
DE	Mouse ischaemic condition related cdna sequence SEQ ID NO:1063.		
KW	Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;		
KW	vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.		
OS	Mus musculus.		
PN	WO200108188-A2.		
PD	22-NOV-2001.		
XX			
PF	18-MAY-2001; 2001WO-JP04192.		
XX			
PR	18-MAY-2000; 2000JP-0145977.		
XX	(UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.		
XX	Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;		

XX	WPI; 2002-034733/04.		
DR	P-PSDB; ABB57374.		
XX			
PT	Examining the ischemic condition (e.g. occlusive ischemia) by measuring		
PT	expression levels of particular genes defined in the specification or		
PT	by determining the expression profile of a gene group comprising these		
XX	genes -		
PS	Claim 2; Page 2680-2685; 2690pp; English.		
XX			
CC	The present invention describes a method for examining ischemic		
CC	conditions, comprising measuring the expression levels of particula		
CC	genes (I) in a test sample or determining the expression profile of a		
CC	gene group in the sample comprising genes selected from (I). The method		
CC	is useful for examining the ischemic condition (e.g. compressive		
CC	ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring		
CC	expression levels of particular genes (ABI99202 to ABI99912, encoding		
CC	the protein sequences in ABB57020 to ABB57374) or by determining the		
CC	expression profile of a gene group comprising these genes. The		
CC	expression levels or expression profiles produced by these genes are		
CC	used as an indicator when screening for ischemic condition-improving		
CC	drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914		
CC	represent PCR primers for a mouse ischaemic condition related sequence,		
CC	which are used in the exemplification of the present invention.		
XX			
SQ	Sequence 2983 BP; 923 A; 634 C; 681 G; 745 T; 0 other;		
Query Match 15.6%; Score 368; DB 24; Length 2983;			
Best Local Similarity 64.9%; Pred. No. 6.7e-87;			
Matches 583; Conservative 0; Mismatches 300; Indels 15; Gaps 2;			
Qy	33	ctctgcagcatgaccaaagaaattctgcagaagttcctgcagatgcagcccaagcaag	92
Db	43	cgcggagagtgaggcaagtgagatcctgcagaggttccatccagaggggtccagccatg	102
Qy	93	aaaattac-----taagaggaggttgcgaatgaattctctgaagctgaaaag	140
Db	103	aagatcgcgatacacaatggggagacaaacttcgcgggacttcgattcgagaga	162
Qy	141	caatctaccagtacaaaggcagacaaaacctatcctacaactgtggtgcgagaatcccaag	200
Db	163	ttgtctaccaaatatagacagaaaagatttatccacagccactgggaaaaagaa	222
Qy	201	aatatcaagaaaaacagataaaggatatatttgcctatgattatagccgggtagaacta	260
Db	223	aatgttataaagaacagataaaggacatactgcatttgatcacagccgagttgaagt	282
Qy	261	tccttgataacctctgatgaggtatccagctacatcaatgccaaacttcattaaaggagt	320
Db	283	actttgaagactccatcccaagattccagatttatataatgcaaatatttattagggtg	342
Qy	321	tatgaccaccaagcttatattccaccagggtctcttatctacaacctctcgacttc	380
Db	343	tatggccaaaaagcatatgtggcaaccccaagggtcttccggaatacacagctatagactc	402
Qy	381	tggaggtgatttgggaatatagttccctatcattgttattggtcatgcattgagatgaa	440
Db	403	tggaggtgatttggaggtataatgtttgatgcgtgattggtgcgtgcgagaattgag	462
Qy	441	atgggaagaaaaagtgtgagcgtactggtggtgcagccagagagatgcagctggaattt	500
Db	463	atgggaagaaaaagtgtgagcgtactggtggtggtgagagatgcatacaattt	522
Qy	501	ggcctttctctgtatcctgtgaaagctgaaaaaagaaatctgattatataatcaggact	560
Db	523	gcaccttttaaaattcttgtgaaatgaacgaagcaacgaactactctccgagaca	582
Qy	561	ctaaaagtttaagttcaaatagtgaaactgaaactatctaccagtttctacagaattgg	620
Db	583	cttttacttgaatttcaaaatgaatcccgctcgctctatcagtttctacgtgaactgg	642

FT		/note= "site of 91 bp intron"	
FT	misc_feature	2532..2545	
FT		/tag= f	
FT		/note= "T-rich sequence required for	
FT		polyadenylation"	
FT	polyA_signal	2740..2747	
FT		/tag= g	
XX			
XX			
PN	WO9748723-A2.		
XX			
PD	24-DEC-1997.		
XX			
XX	17-JUN-1997;	97WO-IB00946.	
PR	19-DEC-1996;	96US-0034286.	
PR	17-JUN-1996;	96US-0019629.	
PR	09-AUG-1996;	96US-0023485.	
PR	13-NOV-1996;	96US-0030860.	
PR	15-NOV-1996;	96US-0030964.	
XX			
XX	(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.		
PA			
XX	Aoki N, Chen Z, Kharitonov AI, Kim YW, Nayler O;		
PI	Ullrich A, Wang HY;		
XX			
DR	WPI; 1998-120302/11.		
DR	P-PSDB; AAW49908.		
XX			
PT	New phosphatase and kinase enzyme(s) - useful in the diagnosis and		
PT	treatment of signal transduction disorders		
XX			
PS	Claim 3; Fig 3a-d; 138pp; English.		
XX			
CC	This cDNA sequence codes for a novel human protein tyrosine		
CC	phosphatase (PTP), designated brain derived phosphatase 1 (BDP-1,		
CC	see AAW49908), that is expressed in most tissues and cell lines at		
CC	basal level, but expressed high in epithelium origin cell lines		
CC	and cancer cell lines. BDP-1 was originally identified in a human		
CC	brain cDNA library, although the full-length clone was isolated		
CC	from the haematopoietic MEG01 cDNA library. The invention relates		
CC	to novel proteins (see AAW49906-14) involved in cellular signal		
CC	transduction and to the nucleic acids (see AAV17097-99) coding for		
CC	them, and provides vectors, host cells, purified recombinant		
CC	proteins, methods for identifying compounds that activate or		
CC	inhibit the novel proteins, as well as methods for the diagnosis		
CC	and treatment of diseases associated with the novel proteins.		
XX			
XX	Sequence 2810 BP; 700 A; 808 C; 808 G; 494 T; 0 other;		
QY	Query Match	8.6%; Score 202.2; DB 19; Length 2810;	
Db	Best Local Similarity	56.5%; Pred. No. 4.7e-43;	
Db	Matches 397; Conservative	0; Mismatches 303; Indels	3; Gaps
QY	199	agaatatcaagaagaaacagatataaggatattttgcccctatgattatagccgggtagaac	258
Db			
QY	207	agaacgtgaggagaacccgctacaagacgtgctgcttatgatcagacgcgagtaatcc	266
Db			
QY	259	tatccctgataaacctctgatgaggattccagctacatacatcaatgccaaacttcattaaggag	318
Db			
Db	267	tctccctgctccagggaaggagacacagcagctacattatggcaacttcctccggcg	326
QY	319	tttatggaccacaaggcttatattgcccaccacgggtcctttatctacaacccctctggact	378
Db			
Db	327	tggatggaacgtggctacattgccaagcgaaggacccttgctcacacccctgctagact	386
QY	379	tcctggaggatgatttgggaatatagtgtccttcattcatgttatggcatgcattggagatg	438
Db			
Db	387	tcctggagactggctgggagtttggggtcaagtgatctctgatggcctgtcagagatag	446
QY	439	aaatcgggaagaaaaagctgtagcgtacttgggctagccagcagagagatgcagctggaat	498
Db			
Db	447	agaatggcgcgaaaggtgtgagcggctactggggccacagcagcagcactgcagactg	506

CC cDNA sequences have been obtained, the full 5' UTR is rarely included
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 205 BP: 63 A: 46 C: 44 G: 52 T: 0 other:
XX

Query Match	6.8%;	Score 160;	DB 21;	Length 205;
Best Local Similarity	100.0%;	Pred. No. 2.1e-32;		
Matches 160;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy 1	tcctccaacctactatagactattttctctgcgactgacacaaagagaattc	60		
Db 46	tcctccaacctactatagactattttctctgcgactgacacaaagagaattc	105		
Qy 61	tgcagaagtctctggatgaggcccaagcagaataattactaaagaggatttgcgaatg	120		
Db 106	tgcagaagtctctggatgaggcccaagcagaataattactaaagaggatttgcgaatg	165		
Qy 121	aattctgaagctgaaaggcnaattctaccaagtacaaggc	160		
Db 166	aattctgaagctgaaaggcnaattctaccaagtacaaggc	205		

RESULT	11	
AAI58418		
ID	AAI58418	standard; cDNA; 4798 BP.
XX		
AC	AAI58418;	
XX		
DT	22-OCT-2001	(first entry)
XX		
DE	Human polynucleotide	SEQ ID NO 621.
XX		
KW	Human;	nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW	peripheral	nervous system; neuropathy; central nervous system; CNS;
KW	Alzheimer's;	Parkinson's disease; Huntington's disease; haemostatic;
KW	amyotrophic	lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW	chemokinetic;	thrombolytic; drug screening; arthritis; inflammation;
KW	leukaemia;	ss.

PS Claim 1; SEQ ID NO 621; 10078pp; English.

xx

CC The invention relates to human nucleic acids (AA157798-AA161369) and

CC the encoded polypeptides (AA338642-AA442213) with neurotropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: Immune system suppression,

CC Activin/Inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and

CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed

CC specification.

xx

SQ Sequence 4798 BP; 1629 A; 850 C; 950 G; 1369 T; 0 other;

Query Match	5.9%;	Score 138.8;	DB 22;	Length 4798;
Best Local Similarity	53.1%;	Pred. No. 3.5e-26;		
Matches 325;	Conservative 0;	Mismatches 277;	Indels 10;	Gaps 1
Qy 164	caaaacatctcctcaaacgtggctggaatgccaaagatcaagaataatcaagaaaaacagatataa	223		
Db 1834	cagcaagtttctctataaaggagcctcgaaagcccttaaccagaataaaaaacgttatgt	1893		
Qy 224	ggatatttggccctatgattagccgggtgagaactatccctgataaacctctgtaggaga	283		
Db 1894	tgacattctctctatgattataaacctggttgaaactctctgagataaacggagagtcag	1953		
Qy 284	ttccagctcatcaatgcacacttcattaaagggagtttatggaccccaaggcttatattgc	343		
Db 1954	gtcaactacataaagtgcagctatatgtaggtttcaagaaccaggaaatacatctgc	2013		
Qy 344	caccagggtcctttatatcaaacctcctggactctctggagtgagtgatttgggaatatag	403		
Db 2014	tgcacaaggtcccagggatgaactgttgatgttctggaggatgtttgggaacagaa	2073		
Qy 404	tgtccttcatcattgttatggcatgcattgagatgaaatgggaagaaaaagtgtagcgc	463		
Db 2074	agccacagttattgtcatgtgtcactcgtatgtagaagaagaaacaggaaacagtgtaga	2133		
Qy 464	ctactgggctgagccaggagagatgcagctggaaatttggcccttctctgtatccctgtga	523		
Db 2134	atatgcccgtcaatggaaggagggcaactcgggcttttgagatgtgtgtgaagaatcaa	2193		
Qy 524	agctgaaaaaaggaaaatctgattatataatcaggagactctaaagttaagttcaatagtga	583		
Db 2194	ccagcacaaaagatgccagattatcatcattcagaaattgaaacattgaaaaaataaaa	2253		
Qy 584	aa-----ctcgaaactatctaccagtttcattacaagaattggccagagaccatgatg	633		
Db 2254	aaaagcaactggaagagaggtgactcacattcagttcaaccagctggccagaccagggg	2313		
Qy 634	taccttcatctatagaccctattctgagctcatctgggatgtaactgtgttaccagaagg	693		
Db 2314	tgcctgaggtcctcaactgtctctcaactgagaaggagagtgatgccttcagcaatt	2373		
Qy 694	atgacagtgttcccatatgcattoactcagtcagtcggtgtgtggaaggagactggtgtattt	753		
Db 2374	tcttcagctgtcccatattggtgcaactcagtcgctggtgtgtggcgacaggaaacctata	2433		
Qy 754	gtgctattgttg 765			
Db 2434	tcggaattgatg 2445			

RESULT 12
AAQ73786


```

XX SQ Sequence 4965 BP; 1683 A; 857 C; 971 G; 1454 T; 0 other;

Query Match          5.5%; Score 129; DB 22; Length 4965;
Best Local Similarity 53.2%; Pred. No. 1.4e-23;
Matches 327; Conservative 0; Mismatches 275; Indels 13; Gaps 2;

QY 164 caaaacctatctacaactgtgctgagatgccagaatatcaagaaaaacagatataa 223
Db 1729 cagcaagttccctataaagaagaagccctttaaccagataataaaacccgttatgt 1788

QY 224 ggatatttgcctatgattatagccgggtagaaactctcctgataacacctctgatgagga 283
Db 1789 tgacattcttcctatgattataaccgtgttgaactctgagataaacggagatgcagg 1848

QY 284 ttccagctacataaagtccaacttcattaaaggagtttatggaccacagcttatattgc 343
Db 1849 gtcaaacatacaaatgccagctattatgtgtttcaagaaccacaggaaatacattgc 1908

QY 344 caccaggggtcctttatctacaaacctcctggactctgagagatgatttggaatatag 403
Db 1909 tgcaaaagtcaccaggatgaactgttgatgtattctctggagatgatttgggaacagaa 1968

QY 404 tgccttatcattgtttatggcatgcagtgagtgatgaatgggaagaaagtgtgagcg 463
Db 1969 agccacagttattgtcatgtgtcactcgatgtgaagaagaaacagaaacaaagttgcaga 2028

QY 464 ctactgggtgagccaggagagatgcagctggaatttggcctttctctgtatcctgtga 523
Db 2029 atactggcgtcaatggaagagggcactcggcgttttggagagtggttgttaaatgct 2088

QY 524 agctgaaaaaaggaaatct----gattataatcaggactctaaaagttaagtccaata 579
Db 2089 aaccagcaaaagatgtcctagatatacattcagaaattggaacattgtaataaaaa 2148

QY 580 gtgaactgaaact-----atcacagtttctacaaagaattggccagaccatg 630
Db 2149 aagaaaaagcaactggaagaggtgactcacattcagttcaccagtgccacagaccg 2208

QY 631 atgtacattcatatagacctattcttgagctatctcgggaatgtaactgtttaccag 690
Db 2209 ggtgctcaggatcctcatctgtcctcctcaaaactgagaaggagagtgaatgccttcagca 2268

QY 691 aggatgacaggttcccatatgcattcactgcagtgctgctgtggaagactggtgtta 750
Db 2269 atttctcagtggtccattgtgtgctcagtgctggtgtgtgtggtggcgacaggaacct 2328

QY 751 tttgtgctattgttg 765
Db 2329 atatcggaattgatg 2343

RESULT 14
AAI60204
ID AAI60204 standard; cDNA; 4968 BP.
XX AC AAI60204;
XX AC AAI60204;
XX AC AAI60204;
DT 22-Oct-2001 (first entry)
XX DE Human polynucleotide SEQ ID NO 4193.
XX DE Human;
XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX KW peripheral nervous system; neuropathy; central nervous system; CNS;
XX KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX KW leukaemia; ss.
XX OS Homo sapiens.
XX PN WO200153312-A1.

```

```

XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US34263.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.
XX PR 03-AUG-2000; 2000US-0653450.
XX PR 14-SEP-2000; 2000US-0662191.
XX PR 19-OCT-2000; 2000US-0693036.
XX PR 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
XX XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
XX P-PSDB; AAM41048.
XX PT Novel nucleic acids and polypeptides, useful for treating disorders
XX PT such as central nervous system injuries -
XX PS Claim 1; SEQ ID NO 4193; 10078pp; English.
XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful
XX CC in gene therapy. A composition containing a polypeptide or polynucleotide
XX CC of the invention may be used to treat diseases of the peripheral nervous
XX CC system, such as peripheral nervous injuries, peripheral neuropathy and
XX CC localised neuropathies and central nervous system diseases, such as
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX CC utilisation of the activities such as: Immune system suppression,
XX CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX CC assays for receptor activity, arthritis and inflammation, leukaemias and
XX CC C.N.S disorders.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification.
XX SQ Sequence 4968 BP; 1678 A; 861 C; 974 G; 1455 T; 0 other;

Query Match          5.5%; Score 129; DB 22; Length 4968;
Best Local Similarity 53.2%; Pred. No. 1.4e-23;
Matches 327; Conservative 0; Mismatches 275; Indels 13; Gaps 2;

QY 164 caaaacctatctacaactgtgctgagatgccagaatatcaagaaaaacagatataa 223
Db 1730 cagcaagttccctataaagaagaagcctcttaaccagataataaaacccgttatgt 1789

QY 224 ggatatttgcctatgattatagccgggtagaaactatcctcgataacacctgatgagga 283
Db 1790 tgacattcttcctatgattataaccgtgttgaactctgagataaacggagatgcagg 1849

QY 284 ttccagctacataaagccaacttcattaggaggagtttatggaccacagcttatattgc 343
Db 1850 gtcaaacatacaaatgccagctatatattgatttccaagaaccacggaaatacattgc 1909

QY 344 caccaggggtcctttatctacaaacctcctcgactcttgagatgatttgggaatatag 403
Db 1910 tgcaaaagtcgccaggatgaaactgttatgtattcttgaggatgatttgggaacagaa 1969

QY 404 tgccttatcattgtttatggcatgcagtgagtgatgaaatgggaagaaaaagtgtgagcg 463
Db 1970 agccacagttattgtcatgtcactcgatgtgaagaagaaacacagaaagtgtgcaga 2029

QY 464 ctactgggtgagccaggagagatgcagctggaatttggccctttctctgtatcctgtga 523

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2002, 10:35:25 ; Search time 54.75 seconds
(without alignments)

308.722 Million cell updates/sec

Title: US-09-600-358A-4

Perfect score: 3615

Sequence: 1 MDQREILQFLDAQSKKIT.....RPSKVKLRSPKSGKNFSL 692

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3543.5	98.0	807	4	US-09-081-345-2
2	2356.5	65.2	802	4	US-09-081-345-18
3	1345.5	37.2	278	4	US-08-821-278A-18
4	1008.5	27.9	272	4	US-08-821-278A-19
5	928.5	25.7	253	2	US-08-685-992-20
6	928.5	25.7	253	2	US-09-144-925-20
7	748	20.7	453	3	US-08-951-260A-7
8	736.5	20.4	453	4	US-08-821-278A-2
9	508	14.1	1337	3	US-08-854-585-2
10	508	14.1	1337	5	PCT-US95-05512-2
11	492.5	13.6	277	2	US-08-685-992-22
12	492.5	13.6	277	2	US-09-144-925-22
13	492.5	13.6	1439	2	US-08-449-644-2
14	492.5	13.6	1439	2	US-08-087-244A-2
15	487	13.5	595	1	US-08-202-389-6
16	485.5	13.4	593	1	US-08-018-129-5
17	485.5	13.4	593	2	US-08-448-250-5
18	485	13.4	1442	1	US-08-015-986A-3
19	485	13.4	1442	2	US-08-446-363-3
20	484	13.4	631	1	US-08-202-389-8
21	479	13.3	1452	2	US-08-652-971-4
22	479	13.3	1452	2	US-08-991-258A-4
23	479	13.3	1452	2	US-08-769-399-4
24	479	13.3	1452	3	US-08-991-953A-4
25	478.5	13.2	1457	2	US-08-652-971-3
26	478.5	13.2	1457	2	US-08-449-644-1
27	478.5	13.2	1457	2	US-08-087-244A-1

Sequence 3, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 12, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 2, Appli
Sequence 12, Appli
Sequence 12, Appli
Sequence 8, Appli
Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-081-345-2
; Sequence 2, Application US/09081345
; Patent No. 6228641
; GENERAL INFORMATION:
; APPLICANT: Bahija Jallal
; APPLICANT: Gregory D. Plowman
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: PTP04 RELATED DISORDERS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" diskette, 1.44 Mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FASTSEQ For Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/081.345
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/047,222
; FILING DATE: May 20, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 234/253
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 807 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-081-345-2

Query Match 98.0%; Score 3543.5; DB 4; Length 807;

```
Best Local Similarity 99.4%; Pred. No. 3e-291;
Matches 691; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 MDQREILQKFLDQAQSKKTKKEEFANFLKLRQSTKYKADKTYPTTVAENAKNIKNRY 60
Db 1 MDQREILQKFLDQAQSKKTKKEEFANFLKLRQSTKYKADKTYPTTVAEKPKNIKNRY 60

QY 61 KDILPYDYSRVLSLTSDESDSYINANFIKGVYGPAYATQGPLSTLLDFWRMIWEY 120
Db 61 KDILPYDYSRVLSLTSDESDSYINANFIKGVYGPAYATQGPLSTLLDFWRMIWEY 120

QY 121 SVLIIVMACHEYEMGKKKCYERYWAEFGEMOLEPFPVSCEAEKRSKDYIIRTLKYKFNS 180
Db 121 SVLIIVMACHEYEMGKKKCYERYWAEFGEMOLEPFPVSCEAEKRSKDYIIRTLKYKFNS 180

QY 181 ETRIIQFYHKNPWDHVPSSIDPILLELWVRCYQEDDSVPICIHCSAGCGRTGVICAI 240
Db 181 ETRIIQFYHKNPWDHVPSSIDPILLELWVRCYQEDDSVPICIHCSAGCGRTGVICAI 240

QY 241 VDYTWMLLDGIIIPENFSVFLIREMRTQPSLVQTOEQYELVYNAVLELFRQMDVIRD 300
Db 241 -DYTWMLLDGIIIPENFSVFLIREMRTQPSLVQTOEQYELVYNAVLELFRQMDVIRD 299

QY 301 KHSGETSOAKHCIPKHNHTLOADSYSPNLPKSTTKAAKMNQORTKMEIKESSDFRFS 360
Db 300 KHSGETSOAKHCIPKHNHTLOADSYSPNLPKSTTKAAKMNQORTKMEIKESSDFRFS 359

QY 361 EISAKELVLHPAKSSTSFDFLELNTSFDKNADTTMKWTKAPPIVGEPLQKHQSIDLGS 420
Db 360 EISAKELVLHPAKSSTSFDFLELNTSFDKNADTTMKWTKAPPIVGEPLQKHQSIDLGS 419

QY 421 LLEFGCSNKPVNAAGRYFNASKVPIIRTKSTPFELIQRETKEVDSKENFSYLESQPHDS 480
Db 420 LLEFGCSNKPVNAAGRYFNASKVPIIRTKSTPFELIQRETKEVDSKENFSYLESQPHDS 479

QY 481 CFYEMQAQKVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGYIYIPLVENYFSS 540
Db 480 CFYEMQAQKVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGYIYIPLVENYFSS 539

QY 541 WPPSGTSSKMSLDLPEKQDGTVPSSLLPTSTSLFSYNSHSSLSLNSPTNISLLNOE 600
Db 540 WPPSGTSSKMSLDLPEKQDGTVPSSLLPTSTSLFSYNSHSSLSLNSPTNISLLNOE 599

QY 601 SAVLATAPRIDDEIPPLPVRTPESTIVVEEAGEFSPNVPKSLSSAVKVKIGTSLWGGT 660
Db 600 SAVLATAPRIDDEIPPLPVRTPESTIVVEEAGEFSPNVPKSLSSAVKVKIGTSLWGGT 659

QY 661 SEPKKFDSDSVILRPSKSVKLRSKPS 685
Db 660 SEPKKFDSDSVILRPSKSVKLRSKPS 684
```

```
RESULT 2
US-09-081-345-18
; Sequence 18, Application US/09081345
; Patent No. 6228641
; GENERAL INFORMATION:
; APPLICANT: Bahija Jallal
; APPLICANT: Gregory D. Plowman
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: PTP04 RELATED DISORDERS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: Storage
```

```
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/081,345
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/047,222
; FILING DATE: May 20, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 234/253
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 802 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-081-345-18
```

```
Query Match 65.2%; Score 2356.5; DB 4; Length 802;
Best Local Similarity 68.5%; Pred. No. 8.5e-191;
Matches 474; Conservative 66; Mismatches 139; Indels 13; Gaps 6;

QY 1 MDQREILQKFLDQAQSKKTKKEEFANFLKLRQSTKYKADKTYPTTVAENAKNIKNRY 60
Db 1 MDQREILQKFLDQAQSKKTKKEEFANFLKLRQSTKYKADKTYPTTVAEKPKNIKNRY 60

QY 61 KDILPYDYSRVLSLTSDESDSYINANFIKGVYGPAYATQGPLSTLLDFWRMIWEY 120
Db 61 KDILPYDYSRVLSLTSDESDSYINANFIKGVYGPAYATQGPLSTLLDFWRMIWEY 120

QY 121 SVLIIVMACHEYEMGKKKCYERYWAEFGEMOLEPFPVSCEAEKRSKDYIIRTLKYKFNS 180
Db 121 SVLIIVMACHEYEMGKKKCYERYWAEFGEMOLEPFPVSCEAEKRSKDYIIRTLKYKFNS 180

QY 181 ETRIIQFYHKNPWDHVPSSIDPILLELWVRCYQEDDSVPICIHCSAGCGRTGVICAI 240
Db 181 ETRIIQFYHKNPWDHVPSSIDPILLELWVRCYQEDDSVPICIHCSAGCGRTGVICAI 239

QY 241 VDYTWMLLDGIIIPENFSVFLIREMRTQPSLVQTOEQYELVYNAVLELFRQMDVIRD 300
Db 240 VDYTWMLLDGIIIPENFSVFLIREMRTQPSLVQTOEQYELVYNAVLELFRQMDVIRD 299

QY 301 KHSGETSOAKHCIPKHNHTLOADSYSPNLPKSTTKAAKMNQORTKMEIKESSDFRFS 357
Db 300 NHLGRIQACSIPEQSLVLEADSCPLDLPKNAMRDVKTTHQHSKQGAESGSSGL 359

QY 358 RTSEISAKEELVLHPAKSSTSFDFLELNTSFDKNADTTMKWTKAPPIVGEPLQKHQSIDL 417
Db 360 RTSTMNAEELVLHSAKSSPFCNLCNCGNNAKAVITRNGQARASPVVGEPLQKYSQSLD 419

QY 418 LGSLLFEGCSNKPVNAAGRYFNASKVPIIRTKSTPFELIQRETKEVDSKENFSYLESQ 477
Db 420 FGSMLFGCSPSALPINTADRYHNSKGPVKRTKSTPFELIQRETNDLAVGDGSCLESQ 479

QY 478 HDSCFV-EMQAQKVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGYIYIPLVENP 536
Db 480 HEHYSURELQVQVAVHVSSEELNYSLP-----CACDASCVPRHSPGALRVHLTSLAEDP 534

QY 537 YFSSWPPSGTSSKMSLDLPEKQDGTVPSSLLPTSTSLFSYNSHSSLSLNSPTNISL 596
Db 535 YFSSSPNSADSCKMSFDLPEKQDGTSPGALLPASSTTFFYSNPHDLSVMTLNTLSFSP 594

QY 597 LNOESAVLATAPRIDDEIPPLPVRTPESTIVVEEAGEFSPNVPKSLSSAVKVKIGTSL 656
```

||||| : : : ||||||| ||||||| || : || : || : ||
Db 595 LNOETAVEAPSRRTDDEIPPLPERTPESFIVVEEAGEPSRVTESL--PLVVTFGASPE 652
QY 657 WGTSEBPKFDDSVILRPSPKSVKLRSPKSGKN 688
Db 653 CSGTSEMKS-HDSVGFPTPSKNVLRSPKSDRH 683
RESULT 3
US-08-821-278A-18
; Sequence 16, Application US/08821278A
; Patent No. 6238902
; GENERAL INFORMATION:
; APPLICANT: Cheng, Jill
; TITLE OF INVENTION: Protein Tyrosine Phosphatases
; FILE REFERENCE: P1010R1
; CURRENT APPLICATION NUMBER: US/08/821,278A
; CURRENT FILING DATE: 1997-03-20
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 18
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo Sapien
US-08-821-278A-18

Query Match 37.2%; Score 1345.5; DB 4; Length 278;
Best Local Similarity 88.5%; Pred. No. 5.6e-106;
Matches 247; Conservative 18; Mismatches 13; Indels 1; Gaps 1;
QY 24 FANFELKLRQSKYKADKTYPTTVAENAKNIKNRYKDLPLDYSRVLSLITSDSDSS 83
Db 1 FASEFLKLRQSKYKADKTYPTTVAQRPNKIKNRKDKLPLDHSVLVSLITSDSDSS 60
QY 84 YINANFIKGYGPKAYIATQGLSTLLDFWRMIWEYSVLIIYMACMEYEMGKKKERYW 143
Db 61 YINASFIKGYGPKAYIATQGLSTLLDFWRMIWEYRILVIVMACMEFEMGKKKERYW 120
QY 144 AEPGEMQLRGPFSVSCAEAKRSDYIIRTLKVKFNSETRTIYOFHYKNPDPHDVPSSID 203
Db 121 AEPGETQLRGPFSISCAEAKRSDYIIRTLKAKFNNETRIIYOFHYKNPDPHDVPSSID 180
QY 204 PILELWDVRCYQEDDSVPICIHCSAGCGRTGVCATVDYTWMLLDGIIIPENFSVFSLI 263
Db 181 PIIQLIWDVRCYQEDDCVPCIHCSAGCGRTGICA-VDTWMLLDGIIIPKFNFSVFNLI 239
QY 264 REMRTQPSLVQTEQYELVYNVAVLELFRQMDVIRDKH 302
Db 240 QEMRTQPSLVQTEQYELVYSVAVLELFRHMDVDSNH 278

RESULT 4
US-08-821-278A-19
; Sequence 19, Application US/08821278A
; Patent No. 6238902
; GENERAL INFORMATION:
; APPLICANT: Cheng, Jill
; TITLE OF INVENTION: Protein Tyrosine Phosphatases
; FILE REFERENCE: P1010R1
; CURRENT APPLICATION NUMBER: US/08/821,278A
; CURRENT FILING DATE: 1997-03-20
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 19
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Homo Sapien
US-08-821-278A-19

Query Match 27.9%; Score 1008.5; DB 4; Length 272;
Best Local Similarity 65.8%; Pred. No. 1.8e-77;

Matches 179; Conservative 46; Mismatches 46; Indels 1; Gaps 1;
QY 24 FANFELKLRQSKYKADKTYPTTVAENAKNIKNRYKDLPLDYSRVLSLITSDSDSS 83
Db 1 FADFMRRLRLSTKYRTEKIYPTATGEKEENAKNRKDKLPLDHSVRKUTLTKTPSQSDS 60
QY 84 YINANFIKGYGPKAYIATQGLSTLLDFWRMIWEYSVLIIYMACMEYEMGKKKERYW 143
Db 61 YINANFIKGYGPKAYVATQGLANTVIDFWRMVWEYVNVVIVMACREFEMGRKKERYW 120
QY 144 AEPGEMQLRGPFSVSCAEAKRSDYIIRTLKVKFNSETRTIYOFHYKNPDPHDVPSSID 203
Db 121 PLYGEDPITFAPEKISCDEQARTDYFIRTLLEFQNESRLYQFHYVNNPDPHDVPSSFD 180
QY 204 PILELWDVRCYQEDDSVPICIHCSAGCGRTGVCATVDYTWMLLDGIIIPENFSVFSLI 263
Db 181 SIIDMISLMRKYQEHEDVPICIHCSAGCGRTGICAL-DYTNWLLKAGKIPEEFNWFNLI 239
QY 264 REMRTQPSLVQTEQYELVYNVAVLELFRQMD 295
Db 240 QEMRTQPSAVQTEQYELVHRAIAQLFEKQL 271

RESULT 5
US-08-685-992-20
; Sequence 20, Application US/08685992
; Patent No. 5912138
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas
; APPLICANT: Flint, Andrew J.
; TITLE OF INVENTION: SUBSTRATE TRAPPING PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESS: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,992
; FILING DATE: 25-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL96-03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 781-861-6240
; TELEFAX: 781-861-9540
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-685-992-20

Query Match 25.7%; Score 928.5; DB 2; Length 253;
Best Local Similarity 65.7%; Pred. No. 9.4e-71;
Matches 167; Conservative 41; Mismatches 43; Indels 3; Gaps 2;

Best Local Similarity 43.1%; Pred. No. 4.5e-55;
Matches 154; Conservative 58; Mismatches 117; Indels 28; Gaps 7;

QY 3 QREILQFLDEAQSCKITKEE-FANFELKLRKSTKYKADKTYPTTVAENAKNKKRYK 61
Db 4 QSLVRSFLEQAGARDHRKGAIIAREFSDIKARSAVWKTGVCSTKAGSQGSKNRYK 63
QY 62 DILPYDYSRVLSLITSDSSYINANFKGVYCPKAYIATQGPLSTLLDFWMIWEYS 121
Db 64 DVVPYDETRVLSLQEGHGDYINANFIRGTDSQAYIATQGPLHTLLDFWRLWFEFG 123
QY 122 VLIIVWACMEYEMGKKCERYWAEPEGMQLEFGPFSVSCAEKR-KSDYIIRTLKVKFNS 180
Db 124 IKVILMACQETENRRKRCERYWAQERE-PLQAGPFCITLTKETALTSIDTLRLTQVTFQK 182
QY 181 ETRTIQFHYKNPDDHVPSSIDPILFELIWDVRCYQEDDSDVPICHSCAGCGRTGVICAI 240
Db 183 ESRVHQLQYMSWPDHGVPSDDHILTMVEEARCLOGLGPGPLCVHCSAGCGRTGVILCA - 241
QY 241 VDTWMLKDGIIIPENFVSFLIREMRTQPSLVQEQYELVYNNAVLELFKQMD---- 296
Db 242 VDVRQLLTQTIPNFSEFVLEVRKQPAVAVQEEQYRFLYHTVAQLFSRTLQNNSP 301
QY 297 -----VTRDKHSCTESQAKHCEIPEKNHTLQADSYS-----PNLPKSTTKA 336
Db 302 LYQNLAKENRAPICKDSSSLTSSA---LPATSRPLGGVLRSSVPGPPTLPMDATYA 355

RESULT 8
US-08-821-278A-2
; Sequence 2, Application US/08821278A
; Patent No. 6238902
; GENERAL INFORMATION:
; APPLICANT: Cheng, Jill
; TITLE OF INVENTION: Protein Tyrosine Phosphatases
; FILE REFERENCE: P1010R1
; CURRENT APPLICATION NUMBER: US/08/821,278A
; CURRENT FILING DATE: 1997-03-20
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 2
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Mus Musculus
US-08-821-278A-2

Query Match 20.4%; Score 736.5; DB 4; Length 453;
Best Local Similarity 44.0%; Pred. No. 4.2e-54;
Matches 144; Conservative 65; Mismatches 109; Indels 9; Gaps 6;

QY 5 EILQKFLDEAQSCKITKEE-FANFELKLRKSTKYKADKTYPTTVAENAKNKKRYKDI 63
Db 6 DLVRSFLEQAGARDHRKGAIIAREFSDIKARSAVWKTGVCSTKAGSRLGNTKKNRYKDV 65
QY 64 LPYDYSRVLSLITSDSSYINANFKGVYCPKAYIATQGPLSTLLDFWMIWEYSVL 123
Db 66 VAIDETRVLSLQEGHGDYINANFIRGTDSQAYIATQGPLHTLLDFWRLWFEFGVK 125
QY 124 IIVWACMEYEMGKKCERYWAEPEGMQLEFGPFSVSCAEKR-KSDYIIRTLKVKFNS 182
Db 126 VILMACQETENRRKRCERYWAQERE-PLKAGPFCITLTKETLNADITLRLTQVTFQKEF 184
QY 183 RTYQFHYKNPDDHVPSSIDPILFELIWDVRCYQEDDSDVPICHSCAGCGRTGVICAI 242
Db 185 RSVHQLQYMSWPDHGVPSDDHILTMVEEARCLOGLGPGPLCVHCSAGCGRTGVILCA -VD 243
QY 243 YTWMLKDGIIIPENFVSFLIREMRTQPSLVQEQYELVYNNAVLELFKQMDVIRDKH 302
Db 244 YVRQLLTQTIPNFSEFVLEVRKQPAVAVQEEQYRFLYHTVAQLFSR---TLQDTS 300
QY 303 SGTESQAKHCIP--EKNHTLQADSYS 327
; Sequence 2, Application PC/TUS9505512

Db 301 PHYQNLKENCAPICKEAFSLRTSSALP 327

RESULT 9
US-08-854-585-2
; Sequence 2, Application US/08854585
; Patent No. 6114140
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas K. and Stman, Arne
; TITLE OF INVENTION: Density Enhanced Protein Tyrosine Phosphatase
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, Suite 6300
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,585
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/237,940
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27866/31954
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1337 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-854-585-2

Query Match 14.1%; Score 508; DB 3; Length 1337;
Best Local Similarity 38.3%; Pred. No. 5.1e-34;
Matches 113; Conservative 51; Mismatches 111; Indels 20; Gaps 7;

QY 15 QSKKITKEEFANFELKLRKST-----KYKADK----TYPTTVAENAKNKKRYKDIPL 65
Db 1017 KSLIRVENEAFYKKQQAADSNCGFAEYEDLKLVGISQPKYAAELAEKNGKNRYNVLP 1076
QY 66 YDYSRVLSLITSDSSYINANFKGVYCPKAYIATQGPLSTLLDFWMIWEYSVLII 125
Db 1077 YDISRVKLSVQTHSTD-DYINANTMPCYHSHKDFIATQGPLNTLKDQFWRWVWKNVYAI 1135
QY 126 VMACMEYEMGKKCERYWAEPEGMQLEFGPFSVSCAEKRKSDYIIRTLKVK--FNSETR 183
Db 1136 IMITKCYEORGTKEEYW--PSKQAQDYGDITVAMTSEIVLPEWITRDTFTVKNIQTSESH 1193
QY 184 TIYQFHYKNPDDHVPSSIDPILFELIWDVRCY--QEDDSVPICHSCAGCGRTGVICAI 241
Db 1194 PLQGFHTSWPDHGVPTDTLLINFRYLVRDYMKQSPESPILVHCSAGVGRIGTFTFID 1253
QY 242 DYTWMLKDGIIIPENFVSFLIREMRTQPSLVQEQYELVYNNAVLELFKQMD 296
Db 1254 RLIVQIENETV---DVYGVIVDLRMHRLPWQTEQDYVFLNOCVLDIVRSQKD 1304

RESULT 10
PCT-US95-05512-2
; Sequence 2, Application PC/TUS9505512

GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas K. and stman, Arne
; TITLE OF INVENTION: Density Enhanced Protein Tyrosine
; NUMBER OF INVENTION: Phosphatase
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 233 South Wacker Drive, Suite 6300
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05512
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27866/31954
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1337 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-05512-2

Query Match 14.1%; Score 508; DB 5; Length 1337;
Best Local Similarity 38.3%; Pred. No. 5.le-34;
Matches 113; Conservative 51; Mismatches 111; Indels 20; Gaps 7;
QY 15 QSKKITKEEFANFLKLRQST-----KYKADK---TYPTTVAENAKNIKKNRYKDILP 65
Db 1017 KSKLRIVENFEAFYKQKQADSCNGFAEYEDLKLGISQPKYAAELAEKNGRNNVLP 1076
QY 66 YDYSRVLSLITSDSSSYINANFIKGYGPKAYIATQGLSTTLDFWRMIWEYSVLII 125
Db 1077 YDISRVKLSVQTHSTD-DYINANYMPGYHSKKDFIATQGLPNTLKFWRWWEKNVYAI 1135
QY 126 VMACMEYEMGKKCERYWAEPMOLEFGFPFVSCEAEKRSYIIRTLKVK--FNSETR 183
Db 1136 IMLTKCQEGRTKCEYW--PSKQADYGDITVAMTSEIVLPEWTRDTFTVKNIQTSSEH 1193
QY 184 TIYQHYKNKPDHVPSSITDPILELIWDVRCY--QEDDSVPICIHCSAGCGRTGVCIAIV 241
Db 1194 PIHQFHTSWPDHGVPTDILLINFRYLVDYMKOSPPSPILVHCSAGVGTGTFFIAID 1253
QY 242 DYTWMLLKDGIIPEFVSFLIREMRTPQSLVQEQYELVYNVAVLEFKRQMD 296
Db 1254 RLIIQIENENTV----DVYGIYVDLMHRPLMVQIEDQYVFLNQCVDIVRSQKD 1304

RESULT 11
US-08-685-992-22
; Sequence 22, Application US/08685992
; Patent No. 5912138
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas
; APPLICANT: Flint, Andrew J.
; TITLE OF INVENTION: SUBSTRATE TRAPPING PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,992
; FILING DATE: 25-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL96-03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 781-861-6240
; TELEFAX: 781-861-9540
; TELEX:
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 277 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-685-992-22

Query Match 13.6%; Score 492.5; DB 2; Length 277;
Best Local Similarity 38.2%; Pred. No. 8.6e-34;
Matches 108; Conservative 40; Mismatches 92; Indels 43; Gaps 8;
QY 45 PITVAENAK--NIKKNRYKDILPYDSRVSELSLITSDSSSYINANFIKGYGPKAYIAT 102
Db 1 PSETSEGDKHHTSKNRYTNILPVNHNTRVOLKIQDKSGSYINANYIDGAY-PROFICT 59
QY 103 QGPLSTTLDFWRMIWEYSVLIIVMACMEYEMGKK---KERYWAEF-----GEMQ 150
Db 60 QGPLNTIADFWRWWEKNRYIIVMLSRSESESESESESESESESESESESESESESESE 119
QY 151 LEFGPFVSCEAEKRSYIIRTLKVKFNSETRTIYQHYKNKPDHVPSSIDPILELIW 210
Db 120 EVFGTYSVELVEVIOCREIITRNIRLTPEGETRDTQYQYEGWPDHNPDPHQPRLAH 179
QY 211 DVRCYQ-----EDDSVPICIHCSAGCGRTGVCIAIV-----DYTWMLLKDG 251
Db 180 SITRNQNIQIPSSDRNVPILVHCSAGVGTGTCTAVIMMKKLDHYFKOLDYNSRI---- 235
QY 252 IIPENFVSFLIREMRTPQSLVQEQYELVYNVAVL-ELFKR 293
Db 236 ----DFNLFISIVLKLREQRPGWVQQLQYLFYCYKILDEIYHR 274

RESULT 12
US-09-144-925-22
; Sequence 22, Application US/09144925
; Patent No. 5951979
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas
; APPLICANT: Flint, Andrew J.
; TITLE OF INVENTION: SUBSTRATE TRAPPING PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.

STREET: Two Militia Drive
 CITY: Lexington
 STATE: MA
 COUNTRY: USA
 ZIP: 02421-4799
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows 95
 SOFTWARE: FastSeq for Windows Version 2.0b
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/144,925
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/685,992
 FILING DATE: July 25, 1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Granahan, Patricia
 REGISTRATION NUMBER: 32,227
 REFERENCE/DOCKET NUMBER: CSHL96-03Z
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 781-861-6240
 TELEFAX: 781-861-9540
 TELEX:
 INFORMATION FOR SEQ ID NO: 22:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 277 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-09-144-925-22

Query Match 13.6%; Score 492.5; DB 2; Length 277;
 Best Local Similarity 38.2%; Pred. No. 8.6e-34;
 Matches 108; Conservative 40; Mismatches 92; Indels 43; Gaps

Qy 45 PTTVAENAK--NIKKNRYKDLIPDYSRVLSLTSDESSYINANFIKGYGPKAYIAT 102
 Db 1 PSETSEGDKKNTSKNRYTLNPNVHTVQLKTKODKSGSYINANYIDGAY-PKQFICT 59
 Qy 103 QGPLSTLLDFWRMIWEYSVLIIIVMACMEYEMGKK---KCRYWAEP-----GEMQ 150
 Db 60 QGPLNTIADFRWVWENRCRIIIVLWSESGENCRKCDRYWPEQIGGQFSGIYNGN 119
 Qy 151 LEFGFVSVCBAEKRSYIIIRTLVKFNSETRTIYQPHYKNWPDHVPSSIDPILLELIW 210
 Db 120 EVFGTYSVELVEVIOCREIITRNIRLTPEGETRDTITQYEGWPDHNPDPHTQPPQLLH 179
 Qy 211 DVRCYQ-----EDDSVPICHCAGCGRTGVCIAV-----DVTWMLKDG 251
 Db 180 SITRQNRQIIPSSDRNPVPIVHCAGVGRGTGCTFCTAVIMMKKLDHYFKQLDYNSRI----- 235
 Qy 252 IIPENFVSFLIREMRTORPSLVOTQEQYELVYNVAVL-ELFKR 293
 Db 236 ----DFNLFSLVLRQRPQWVOOLEQYLFICYTKTILDEIYHR 274

RESULT 13
 US-08-449-644-2
 ; Sequence 2, Application US/08449644
 ; Patent No. 5856162
 ; GENERAL INFORMATION:
 ; APPLICANT: Schlössinger, Joseph
 ; APPLICANT: Sap, Jan M.
 ; APPLICANT: Ulrich, Axel
 ; APPLICANT: Vogel, Wolfgang
 ; APPLICANT: Fuchs, Miriam
 ; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
 ; TITLE OF INVENTION: PHOSPHATASE-KAPPA
 ; NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:
 ADDRESSEE: PENNIE & EDMONDS
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/449,644
 FILING DATE: 24-MAY-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/087,244
 FILING DATE: 01-JUL-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Coruzzi, Laura A.
 REGISTRATION NUMBER: 30,742
 REFERENCE/DOCKET NUMBER: 7683-042
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-790-9090
 TELEFAX: 212-869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1439 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-449-644-2

Query Match 13.6%; Score 492.5; DB 2; Length 1439;
 Best Local Similarity 29.1%; Pred. No. 1.2e-32;
 Matches 136; Conservative 77; Mismatches 167; Indels 87; Gaps 16;

Qy 21 KEFANEFLKLRSTKYKADKTYPTVAENAKNIKKNRYKDLIPDYSRVLSLTSDE 80
 Db 888 KEYESFP---EGOSASW-----DVAKKDONRAKNRYGNIITAYDHSRVLQPVEDDP 936
 Qy 81 DSSYINANFIKGYGPKAYIATQGPLSTLLDFWRMIWEYSVLIIIVMACMEYEMGKKKE 140
 Db 937 SSDYINANYIDGQRPISHYIATQGPVHETVYDFWRMIWQEOSACIVWNTLVNVEGRVKY 996
 Qy 141 RYWAEPGEMQLEFGPFSVSCBAEKRSYIIIR--TLKVKFNSETRTIYQPHYKNWPDHVD 198
 Db 997 KYWPDDEV---YGDVKVTCVEMEPLEAYVVVTRTFLERRGYNEIREVKQFHFTGWPDHGV 1053
 Qy 199 PSSIDPILLELIWVRCYOEDDSVPICHCAGCGRTGVCIAIVDTWMLLKDGIIPEPNS 258
 Db 1054 PYHATGLLSFRRVKLSNPPSAGPIVHCAGAGRTGCIIVIDIMLDMAREGVV-----D 1109
 Qy 259 VFSILIREMRTORPSLVOTQEQYELVYNVAVL-----FK-RQMDVIR----- 299
 Db 1110 IYNCVKALRSRRINNVQTEQYIFIHDAILEACLCGETAIPVCEFKAAAYFDMIRIDSOTN 1169
 Qy 300 DKHSGTESQAKH-----CIPEKNHTLQ--ADSVSPN--LPKSTTKAAKMMN 341
 Db 1170 SSKLKDEFQTLNSVTPRLQAEDCSLACLP-RNHDKNRNMDMLPPDCLPFLITIDGESSN 1228
 Qy 342 QORTKM--EIKESSSDFRTSEISAKHEELVLHPAKSSTSFDFLELNYSF----- 388
 Db 1229 YINAALMDSYRQPAAF-----IVTQYPLNPKDFWRLVYDYGCTSVMLNEV 1276
 Qy 389 DKNADTTMKWQTKAF----PIVGEPLQHQSLDGLSLLFEGCSNSKP 431
 Db 1277 DLSQGCQPYWPEEGLMRYGPIQVECMSCMDCDVINFRICNLTRP 1323

RESULT 14
US-08-087-244A-2
; Sequence 2, Application US/08087244A
; Patent No. 5863755
; GENERAL INFORMATION:
; APPLICANT: Schlensing, Joseph
; APPLICANT: Sap, Jan M.
; APPLICANT: Ullrich, Axel
; APPLICANT: Vogel, Wolfgang
; APPLICANT: Fuchs, Miriam
; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASE-KAPPA
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,244A
; FILING DATE: 01-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-042
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1439 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-087-244A-2

Query Match 13.6%; Score 492.5; DB 2; Length 1439;
Best Local Similarity 29.1%; Pred. No. 1.2e-32;
Matches 136; Conservative 77; Mismatches 167; Indels 87; Gaps 16;
QY 21 KEFANEFLKQSTKYKADKTYPTVAENAKNIKKNRYKIDILPYDSRVLSLTSDE 80
DB 888 KEEYESFF---EQGSASW-----DVAKKQNRKRNKGNIIADHSVILQPVDDP 936
QY 81 DSSYINANFTKYGKAYATGGPLSTLLDFWRMIWEYSVLIIVMACMEYEMGKKCE 140
DB 937 SSDYINANYIDGQPSHYLATQPVHETVYDFWRMIWEQSACIVMTNLNVEGVKCY 996
QY 141 RYWAEPCEMQLGEPFSVSCAEKRSYIIR--TLKVFNSETRIYQPHYKNWPDHV 198
DB 997 KYPDPDTEV---YGFDFKVTCEVNEPLAEYVVRTTLERRGYNEIREVKQPHFTGWDHGV 1053
QY 199 PSSIDPTELEIWDVRCYQEDDSVPICHTCSAGCGRTGCAIVDYTMLLKDGIIPEFS 258
DB 1054 PYHATGLLSFIRRVKLSNPSPAGPIVHCSSAGAGRTGCIIVIMLDMAREGVV----D 1109
QY 259 VFSLIREMTRQPSLVQTEQYELVYNVAVLEL-----FK-RQMDVIR----- 299
DB 1110 IYNCVKALRSRRINMYQTEQYIFIHDAILEACLCGETAIPVCEFAAAYFDMIRIDSQTN 1169
QY 300 DKHSCTESQAKH-----CIPEKNHTLQ--ADSYSPN--LPKSITKAAKMMN 341
DB 1170 SSHLKDEFQTLNSVTPRLQAEDCSIACLP--RNHDKNRNFMPLPPDRCLPFLITIDGESSN 1228

QY 342 QORTKM--EIKESSSDFRTSEISAKBELVHPAKSSTSPDFLELNYSF----- 388
DB 1229 YINAALMDSYRQPAAF-----ITYQYPLPNTVDFWRLYDYCYCTSIYMLNEV 1276
QY 389 DKNADTTMKWTKAF-----PIVGEPLQKHSLDGLSLFEGCSNSKP 431
DB 1277 DLSQGCPOQYWPPEGMLRYGPIQVECMSCSMDCDVINRIFRICNLTRP 1323
RESULT 15
US-08-202-389-6
; Sequence 6, Application US/08202389
; Patent No. 5536636
; GENERAL INFORMATION:
; APPLICANT: Freeman Jr., Robert M.
; APPLICANT: Plutzky, Jorge
; APPLICANT: Neel, Benjamin G.
; APPLICANT: Rosenberg, Robert D.
; TITLE OF INVENTION: IDENTIFICATION OF NOVEL TYROSINE
; TITLE OF INVENTION: PHOSPHATASES HAVING SH2 DOMAINS
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/202,389
; FILING DATE: 28-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/983,926
; FILING DATE: 01-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/829,141
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/721,112
; FILING DATE: 26-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: BIH92-05MA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 595 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-202-389-6

Query Match 13.5%; Score 487; DB 1; Length 595;
Best Local Similarity 33.0%; Pred. No. 8.4e-33;
Matches 128; Conservative 69; Mismatches 129; Indels 62; Gaps 14;
QY 15 QSKKITKEEFANEFLLKROSTKYKADKTYPTVAENAKNIKKNRYKIDILPYDSRVLS 74
DB 235 ESDTAKAGFEWEEFSLQKEVK-----NLHQLEGORPENKGNRYKNILPFDHSRVILQ 290
QY 75 LITSDEDSYINANFTKGVYGP-----KAYIATQGPLSTLLDFWRMIWEYSVLIIYMA 128

Db 291 GRDSNIPGSDYINANYIKNQLLGPDENAKTYIASQGCLEATVNDFWQAMQENSRIYMT 350
Qy 129 CMYEMGKKKCYWAEPEGMQLEFGPFSVSCAEKRSQDYIIRTLKVK--FNSE-TRTI 185
Db 351 TREVEKGRNKCVPWPEVG-MQRAYGPYSVTNVGEHDTTEYKRLTLQVSPDLNDGLIREI 409
Qy 186 YOFHYKNMPDHDVPSSIDPIILELIWDVRCYQED--DSVPICIHCSAGCGRTGVICAIVDY 243
Db 410 WHYQYLSWPDHGVSEPGGVLSFLDQINORQESLPHAGPIIVHCSAGIGRTGTIIVI--- 466
Qy 244 TWMLLKDGIIIPENFSVSL-----IREMRTORPSLVOTQOYELVYNV---LELF 291
Db 467 -----DMLMENISTRGLDCDIDIQKTQIMVRAQRSGMVQTEAQYKFIYVAIAQFIETT 519
Qy 292 KROMDVIRDKHSGTESQAKHCIPKKNHTLOADSYSNLPKSTTKAAK-----M 339
Db 520 KKKLEVIQSQ-KGQSEYGN-----TTPPAKNNAHAKASRTSSKHKEDVYENL 567
Qy 340 MNQORTKMEIKESSFDFRTSEISAKEE 367
Db 568 HTKNKREKVKQRSADKESKSGSLKRR 595

Search completed: May 26, 2002, 12:17:53
Job time: 6148 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2002, 10:34:06 ; Search time 119.16 Seconds
(without alignments)
645.041 Million cell updates/sec

Title: US-09-600-358a-4
Perfect score: 3615
Sequence: 1 MDQREILQKFLDEAQSCKIT.....RPSKSVKLKSPKSGKNFWSL 692

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3615	100.0	692	AA128653	Human Cytoplasmic
2	3573	98.8	808	AA128652	Human Cytoplasmic
3	3543.5	98.0	807	AA189247	Human PTP04. Homo
4	2971.5	82.2	799	AA18623	Human tyrosine pho
5	2356.5	65.2	802	AA128654	Murine Cytoplasmic
6	1158	32.0	780	AA151205	Human PTP-PEST Gen
7	1157	32.0	780	AA151201	Human protein tyro
8	1150	31.8	780	AA151201	Human PTP-PEST mut
9	1148	31.8	780	AA151202	Human PTP-PEST mut
10	1147	31.7	780	AA151202	Human mutant PTP-P
11	1127.5	31.2	775	AA167250	Mouse protein tyro

12	1117.5	30.9	775	21	AA167251	Mutant mouse PTP-P
13	1058	29.3	331	21	AA158898	Breast and ovarian
14	933	25.8	312	22	AA178281	Human PTP-PEST. H
15	921.5	25.5	253	22	AA159383	Human protein tyro
16	748	20.7	433	19	AA149906	Rat protein tyrosi
17	736.5	20.4	453	18	AA137254	Novel non-receptor
18	708	19.6	458	19	AA149908	Human brain derive
19	509	14.1	853	19	AA185031	Protein tyrosine p
20	508	14.1	1337	16	AA185203	huDPP-1. Homo sap
21	502	13.9	1149	22	AA141048	Human polypeptide
22	496	13.7	1440	20	AA129591	Human protein phos
23	493	13.6	274	22	AA159385	Protein tyrosine p
24	492.5	13.6	341	22	AA178283	PTP1. Unidentifie
25	492.5	13.6	1439	15	AA163632	Human SH-PTP1 (pro
26	487	13.5	595	17	AA199312	Human SH-PTP1 (pro
27	486.5	13.5	593	15	AA152991	Human protein-tyro
28	485	13.4	595	22	AA159216	SHP-1 protein. Un
29	485	13.4	595	22	AA159232	SHP-1 activated tr
30	485	13.4	595	22	AA159233	SHP-1 activated tr
31	485	13.4	595	22	AA159238	SHP-1 D59A mutant
32	485	13.4	595	22	AA159239	SHP-1 E74A mutant
33	485	13.4	597	21	AA126873	Human Src homolog
34	485	13.4	621	21	AA152289	Human HCP phosphod
35	485	13.4	824	21	AA152288	Human HCP/GST fusi
36	485	13.4	843	19	AA185030	Green fluorescent
37	477	13.2	595	22	AA159228	SHP-1 activated do
38	477	13.2	595	22	AA159231	SHP-1 activated do
39	477	13.2	595	22	AA159235	SHP-1 activated qu
40	477	13.2	595	22	AA159237	SHP-1 activated qu
41	477	13.2	623	22	AA125675	Human protein sequ
42	477	13.2	1452	22	AA179159	Human protein SEQ
43	477	13.2	1455	22	AA180143	Human protein SEQ
44	476	13.2	1445	15	AA158809	Human RPTP-gamma.
45	475.5	13.2	1237	19	AA144729	Chicken protein ty

ALIGNMENTS

RESULT 1	
AA128653	
ID	AA128653 standard; Protein; 692 AA.
XX	
AC	AA128653;
XX	
DT	01-OCT-1999 (first entry)
XX	
DE	Human Cytoplasmic phosphatase, Lyp2 protein.
XX	
KW	Lymphoid Protein Tyrosine Phosphatase; Lyp protein; immunosuppressant;
KW	intracellular tyrosine phosphatase; PTPase; fetal liver; transplanti;
KW	resting lymphoid cell; protein tyrosine kinase; PTKs; lymphocyte;
KW	T cell antigen receptor signalling; cytokine receptor signalling;
KW	autoimmune disease; intronic sequence; alternative mRNA splicing.
XX	
OS	Homo sapiens.
XX	
PH	Key
FT	Domain
FT	Location/Qualifiers
FT	27..289
FT	/label= PTPase_domain
FT	/note= "Single catalytic protein tyrosine phosphatase
FT	domain"
FT	469..472
FT	/label= NXXY_motif
FT	/note= "Unique sequence recognised by phosphotyrosine
FT	binding (PTB) domain"
FT	615..623
FT	/label= SH3_binding_site
FT	/note= "Proline rich sequence"
XX	
PN	WO9936548-A1.
XX	
PD	22-JUL-1999.

XX PF 18-JAN-1999; 99WO-CA00038.
XX PR 16-JAN-1998; 98CA-2220853.
XX PA (HSCR-) HSC RES & DEV LP.
XX PI Roifman CM;
XX WPI: 1999-444404/37.
XX DR N-PSDB; AAX90696.
XX PT New nucleic acid encoding intracellular tyrosine phosphatase and related proteins, used to modulate signalling through T cells, particularly as immunosuppressant
XX PS Claim 4b; Page 55; 105pp; English.
XX CC The present protein sequence is that of the cytoplasmic, Lymphoid Protein Tyrosine Phosphatase, Lyp2 protein that has a single catalytic domain. The non-catalytic portion of the phosphatase contains unique sequences, including a single PEST sequence rich in Pro, Glu or Asp, Ser and Thr. It is expressed significantly in fetal liver and in resting lymphoid cells. Lyp2 is an isoform of the Lyp1 gene that arises by alternative splicing of the mRNA. The intronic sequence of Lyp1 encodes for the C-terminal 7 aminoacids and part of the 3'/untranslated region of Lyp2. Lyp2 shares sequence identity with the murine phosphatase 270PEP. Lyp proteins are important for regulation of T cell antigen and cytokine receptor signalling and for early and late stages of T cell differentiation. Lyp2 has immunosuppressive activity. Compounds that increase expression of Lyp protein can be used as immunosuppressive agents to reduce or prevent T cell activation or proliferation, to control thymocyte differentiation, to treat autoimmune diseases and transplant situations.
XX SQ Sequence 692 AA;
Query Match 100.0%; Score 3615; DB 20; Length 692;
Best Local Similarity 100.0%; Pred. No. 8.2e-300;
Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDQREILQKFLDRAQSKKIKTEFANFELKLRQSTKYKADKTYPTTVAENAKNIKKNRY 60
DB 1 mdgreilqlkfldeaqskkikteefanefllkkrqstkykadktypttvaenaknikknry 60
QY 61 KDILPYDSRVELSLTSDSSYINANFKGVYGPAYATOGPLSTLLDFWRMWEY 120
DB 61 kdilpydsrvelsltsdssyinanfllkgvygpayatogplstlldfwrmlwey 120
QY 121 SVLIIVNACMEYEMGKKKCYWAEPEGMQLERGFPSVSCAEKRKSDYIIRTLKVKFNS 180
DB 121 svliivnacmeyemgkkkcywaepegmqlerfgpsvscaeakrksdyiirtlkvkfn 180
QY 181 ETRTIYOFHYKNPDPHDVPSIDPILEIWDVRCYQEDDSVPICIHCSAGCGRTGVICAI 240
DB 181 etrtiyofhyknpdhdvpsidpileiwdvrcyqeddsvpicihcsagcgrtgvicai 240
QY 241 VDYTWMLLKDGIIPENFSVFSLSIREMTORPSLVQTQEQYELVYNVAVLELFKQMDVIRD 300
DB 241 vdytwmlldkgilpenfsvflsiremtorpslvqtqeqyelvynvavlelfkrqmdvird 300
QY 301 KHSGETSQAKHCIPKHNHTLQADSYPNLKSTTKAAKMMNQORTKMEIKRESSSFDPRTS 360
DB 301 khsgetsqakhcipekhnhtlqadsysnlpksttkaaakmmnqortkmeikesssfdfrts 360
QY 361 EISAKEELVLPKAKSSYFDFLELUNYFSDKNADTTMKQTKAFPIVGEPIQKHQSLLDLS 420
DB 361 eisakeelvlpkakssyfdflelnysfdknadtmtkqtkafpivgepiqkhqslldls 420
QY 421 LLFEGCSNKPVNAGRYFNYSKVPITRTKTPPELIQOQRETKEVDKSNFYLESQPHDS 480
DB 421 llfegcsnkpvnagryfnyskvpitrtktppeliqoqretkevdkensfylesqphds 480

QY 481 CFVEMOAKVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGYSYIPLVENPYFSS 540
DB 481 cfvemgaqkvmhvssaelnyslpydskhqirnasnvkhhdsalgvysyiplvenpyfss 540
QY 541 WPPSGTSSKMSIDLPEKQDGTVPFSSLLPTSTSLFSYNNSSHSSLSLNSPTNMISSLLNQE 600
DB 541 wppsgtsskmsidlpekqdggtvpssllptstslfsyynsshsslslnsptnmissllnqe 600
QY 601 SAVLATAPRIDDEIPPLPVRTPEFIVVEAGEESPVPKSLSSAVKVKICTSLEWGGT 660
DB 601 savlatapriddeipplpvrtpesfivveageespnvpkslssavkvkigtslewggt 660
QY 661 SBPKKFDSDSVILRPSKSVKLRSPKSGKNFSWL 692
DB 661 sepkkfdsdsvilrpsksvklrpsksgknfswl 692
RESULT 2
AAY28652
ID AAY28652 standard; Protein; 808 AA.
XX AC AAY28652;
XX DT 01-OCT-1999 (first entry)
XX DE Human Cytoplasmic phosphatase, Lyp1 protein.
XX KW Lymphoid Protein Tyrosine Phosphatase; Lyp protein; Lymphoid cell;
KW Intracellular tyrosine phosphatase; PPase; lymphocyte; thymocyte;
KW T cell; B cell; protein tyrosine kinase; PTKs; immunosuppressant;
KW T cell antigen receptor signalling; cytokine receptor signalling;
KW autoimmune disease; transplant.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Domain 27..289 /label= PPase_domain
FT /note= "Single catalytic protein tyrosine phosphatase domain"
FT Domain 469..472 /label= NXXY_motif
FT /note= "Unique sequence recognised by phosphotyrosine binding (PTB) domain"
FT Binding-site 615..623 /label= SH3_binding_site
FT /note= "Proline rich sequence"
FT Binding-site 694..701 /label= SH3_binding_site
FT /note= "SH3 binding site"
FT Region 702..736 /label= PEST_sequence
FT /note= "Sequence rich in Pro, Glu or Asp, Ser and Thr"
FT Region 741..745 /label= PEST_sequence
FT /note= "Consensus sequence recognised by p34cdc2 kinase"
FT Binding-site 768..772 /label= SH3_binding_site
FT /note= "Proline rich sequence"
FT Binding-site 796..804 /label= SH3_binding_site
FT /note= "Proline rich sequence"
XX WO9936548-A1.
XX PN 22-JUL-1999.
XX PD 18-JAN-1999; 99WO-CA00038.
XX PF 16-JAN-1998; 98CA-2220853.
XX PR (HSCR-) HSC RES & DEV LP.
XX PA


```
Query Match      98.0%; Score 3543.5; DB 20; Length 807;
Best Local Similarity 99.4%; Pred. No. 1.4e-293;
Matches 681; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 MDQREILQKFLDEAQSCKITKEEFANEFUKLRQSTKYKADKTYPTTVAENAKNIKKNRY 60
DB 1 mdqreilqkfldeagskkltkeefaneflklrqstkykadktypttvaekpknikknry 60
QY 61 KDILPYDYSRVLSITSDSSYINANFIKGVYGPKAVIATQGPLSTLLDFWRMIWEY 120
DB 61 kdilpydysrvlsitdsdssyinanfikvgypkaviatqgplstllcldfwrmiwey 120
QY 121 SVLIIVMACMEYEMGKKKERYWAEPGEMQLEFGPFSVCEAEKRSKSDYIIRTLVKVFN 180
DB 121 svliivmacmeyemgkkkcerywaepgemqlefgpfsvsceaeakrksdyiirtlvkvn 180
QY 181 ETRTIYQFHYKNWPDHVDVSSIDPILLELWDVRCYQEDDSVPICHCAGCGRTGVICAI 240
DB 181 etrtiyqfhyknwphdvpssidpillelwdvrcyqeddsvpicchcsagcgrtgvicai 240
QY 241 VDYTWMLLKDGIIPENFVSFSLIREMRTQPSLVOTQOEYELVYNVAVLELFRQMDVIRD 300
DB 241 -dytwmlldgiiipenfsvsliremrtqpslvotqoeylvynvavlelfrqmdvird 300
QY 301 KHSGETSQAKHCIPKHNHTLQADSYSPNLPKSTTTAAKMMNQORTKMEIKESSSDFRTS 360
DB 300 khsgetsqakhcipeknhtlqadsyspnlpksttkaakmmnqqrktmeikesdfrts 360
QY 361 EISAKRELVLHPAKSSTSFDFLELNYSPDKNADTTMKWOTKAPPIVGEPLQKHSIDLGS 420
DB 360 eisakeelvlhpaksstsfdflelnyspdknadtmkwotkappivgeplqkhsidlgs 420
QY 421 LLFEGCSNKPVNAAGRYNSKVPIRTKSTPELIIQORETKEVDSKENFVLESOPHDS 480
DB 420 llfegcsnkpvnagrynskvpiirtkstpfelliqoretkevdskenfylesophs 480
QY 481 CFVEMOQAKVMHVSSAELNYSILPYDSKHQIRNASNVKHHDSALGVYSYIPLVENPYFSS 540
DB 480 cfvemqakvmhvssaelnysilpydskhqirnasnvkhhdsaalgvysyiplvenpyfss 540
QY 541 WPPSGTSSKMSLDLPEKQDGTVPSSLLPTSTSLFSYNNSSLSLNSPTNISSLNQE 600
DB 540 wppsgtsskmsldlpekqdggtvfpsllptstslfsynnsslslnsptnisslnqe 600
QY 601 SAVLATAPRIDDEIPPLPVRTPEFIVVEEAGEFSPNPKSLSSAVKVKIGTSLWGGT 660
DB 600 savlatapriddeipplpvrtpesfivveeagefsonpvpkslssavkvkigtswgg 660
QY 661 SEPKKFDSDSVILRPSKSVKLRSPKS 685
DB 660 sepkkfddsvilrpsksvklrspks 684
```

```
RESULT 4
AAG78623
ID AAG78623 standard; Protein; 799 AA.
AC AAG78623;
XX
XX
XX
XX
DT 04-DEC-2001 (first entry)
XX
DE Human tyrosine phosphatase.
XX
XX Human; tyrosine phosphatase; hTTP.
XX
OS Homo sapiens.
XX
XX CN1302899-A.
XX
PD 11-JUL-2001.
XX
XX 29-OCT-1999; 99CN-0119935.
```

```
XX 29-OCT-1999; 99CN-0119935.
XX (NANF-) NANFANG RES CENT STATE HUMAN GENE GROUP.
XX
XX Ren S, Wu T, Qian B;
XX
XX WPI; 2001-550591/62.
XX N-PSDB; AAH79342.
XX
XX Human protein tyrosine phosphatase and its coding sequence -
XX
XX Claim 4; Page 12-13(Disclosure); 29pp; Chinese.
XX
XX The present invention provides the protein and coding sequences of human
XX tyrosine phosphatase hTTP. The protein is expressed in human normal
XX suprarenal tissue. The present sequence is the protein of the invention.
XX
XX Sequence 799 AA;

Query Match      82.2%; Score 2971.5; DB 22; Length 799;
Best Local Similarity 87.3%; Pred. No. 1e-244;
Matches 598; Conservative 17; Mismatches 61; Indels 9; Gaps 9;

QY 1 MDQREILQKFLDEAQSCKITKEEFANEFUKLRQSTKYKADKTYPTTVAENAKNIKKNRY 60
DB 1 mdqreilqkfldeagskkltkeefaneflklrqstkykadktypttvaekpknikknry 60
QY 61 KDILPYDYSRVLSITSDSSYINANFIKGVYGPKAVIATQGPLSTLLDFWRMIWEY 120
DB 61 kdilpydysrvlsitdsdssyinanfikvgypkaviatqgplstllcldfwrmiwey 120
QY 121 SVLIIVMACMEYEMGKKKERYWAEPGEMQLEFGPFSVCEAEKRSKSDYIIRTLVKVFN 180
DB 121 svliivmacmeyemgkkkcerywaepgemqlefgpfsvsceaeakrksdyiirtlvkvn 180
QY 181 ETRTIYQFHYKNWPDHVDVSSIDPILLELWDVRCYQEDDSVPICHCAGCGRTGVICAI 240
DB 181 etrtiyqfhyknwphdvpssidpillelwdvrcyqeddsvpicchcsagcgrtgvicai 240
QY 241 VDYTWMLLKDGIIPENFVSFSLIREMRTQPSLVOTQOEYELVYNVAVLELFRQMDVIRD 300
DB 241 -dytwmlldgiiipenfsvsliremrtqpslvotqoeylvynvavlelfrqmdvird 300
QY 301 KHSGETSQAKHCIPKHNHTLQADSYSPNLPKSTTTAAKMMNQORTKMEIKESSSDFRTS 360
DB 300 khsgetsqakhcipeknhtlqadsyspnlpksttkaakmmnqqrktmeikesdfrts 360
QY 361 EISAKRELVLHPAKSSTSFDFLELNYSPDKNADTTMKWOTKAPPIVGEPLQKHSIDLGS 420
DB 360 eisakeelvlhpaksstsfdflelnyspdknadtmkwotkappivgeplqkhsidlgs 420
QY 421 LLFEGCSNKPVNAAGRYNSKVPIRTKSTPELIIQORETKEVDSKENFVLESOPHDS 480
DB 417 lcl-gmsnslv-csrkifnrcqyp-dqinsfeliqq-egdgdsdkktflf-esqphs 471
QY 481 CFVEMOQAKVMHVSSAELNYSILPYDSKHQIRNASNVKHHDSALGVYSYIPLVENPYFSS 540
DB 472 cfvemqakvmhvssaelnysilpydskhqirnasnvkhhdsaalgvysyiplvenpyfss 531
QY 541 WPPSGTSSKMSLDLPEKQDGTVPSSLLPTSTSLFSYNNSSLSLNSPTNISSLNQE 600
DB 532 wppsgtsskmsldlpekqdggtvfpsllptstslfsynnsslslnsptnissrieqe 591
QY 601 SAVLATAPRIDDEIPPLPVRTPEFIVVEEAGEFSPNPKSLSSAVKVKIGTSLWGGT 660
DB 592 savlatapriddeiplhlygtpesfivveeagefsonpvpkslssavkvkigtswgg 651
QY 661 SEPKKFDSDSVILRPSKSVKLRSPKS 685
DB 652 sepkkfddsvilrpsksvklrspks 676
```

```

RESULT      5
AAAY28654
ID  AAAY28654 standard; Protein; 802 AA.
XX
AC  AAAY28654;
XX
DT  01-OCT-1999 (first entry)
XX
DE  Murine Cytoplasmic phosphatase, Z70PEP protein.
XX
XX  Lymphoid Protein Tyrosine Phosphatase; Lyp protein; lymphoid cell;
KW  intracellular tyrosine phosphatase; PTPase; lymphocyte; murine;
KW  protein tyrosine kinase; PTKs; immunosuppressant; PEST sequence;
KW  T cell antigen receptor signalling; autoimmune disease; transplant;
KW  cytokine receptor signalling.
XX
OS  Mus sp.
FH
FH  Key
FH  Domain
FT  Location/Qualifiers
FT  27..288
FT  /label= "PTase domain"
FT  /note= "Catalytic protein tyrosine phosphatase domain"
FT  Binding-site
FT  613..621
FT  /label= "SH3 binding site"
FT  /note= "Proline rich sequence"
FT  Binding-site
FT  689..695
FT  /label= "SH3 binding site"
FT  /note= "Proline rich sequence"
FT  Binding-site
FT  790..798
FT  /label= "SH3 binding site"
FT  /note= "Proline rich sequence"
XX
XX  WO9936548-Al.
XX
XX  22-JUL-1999.
XX
XX  18-JAN-1999; 99WO-CA00038.
XX
XX  16-JAN-1998; 98CA-2220853.
XX
XX  (HSCR-) HSC RES & DEV LP.
XX
XX  Roifman CM;
XX
XX  WPI; 1999-444404/37.
XX
XX  New nucleic acid encoding intracellular tyrosine phosphatase and
XX  related proteins, used to modulate signalling through T cells,
XX  particularly as immunosuppressant
XX
XX  Disclosure; Page 63-64; 105pp; English.
XX
XX  The present protein sequence is that of the murine phosphatase, Z70PEP
XX  that has a single catalytic domain. The non-catalytic portion of the
XX  phosphatase contains unique sequences, including five PEST sequences
XX  rich in pro, Glu or Asp. Ser and Thr. Z70PEP shares about 70% sequence
XX  identity with the human cytoplasmic phosphatase Lyp1. Lyp proteins are
XX  important for regulation of T cell antigen and cytokine receptor
XX  signalling and for early and late stages of T cell differentiation.
XX  Z70PEP has immunosuppressive activity. Compounds that increase
XX  expression of Lyp protein can be used as immunosuppressive agents to
XX  reduce or prevent T cell activation or proliferation, to control
XX  thymocyte differentiation, to treat autoimmune diseases and transplant
XX  situations.
XX
XX  Sequence 802 AA;
XX
XX  Query Match 65.2%; Score 2356.5; DB 20; Length 802;
XX  Best Local Similarity 68.5%; Pred. No. 3.9e-192;
XX  Matches 474; Conservative 66; Mismatches 139; Indels 13; Gaps 6;

```

```

QY 1 MDQREILQKFLDEAQAQKKITKEEFANEFKLKRQSTKYKADKTYPTTVAENAKNIKNRY 60
DB 1 mdqreilqlqlkeaqkklnseefaseflklkrqstkykadiptttvaqrpnknkny 60
QY 61 KDILPYDYSRVLSLITSDSSYINANFIKGYGPKAYIATQGPLSTTLLDFWRMWEY 120
DB 61 kdilpydhsiveislitsdedssyinasfikgygpkayiatqgplsttlldfwrmlwey 120
QY 121 SVLIIVMACMEYEMGKKCERYWAEPGEMOLEFGPPFSVSCAEKRSYDIIRTLKVFNS 180
DB 121 rliivmacemefengkkcerywaepgetqlqfpgfsisceaeakksdyirtlkakfnn 180
QY 181 ETRTIYQFHYKNWPDHDPSSIDPILLEIWDVRCYQEDDSVPICIHGACGRTGVICAI 240
DB 181 etriiyqfhyknwpdhdpssidpilhqlwdmrcygeddcvpcihcsagcgrtgvisa- 239
QY 241 VDYTWMLLKDGIIIPENFVSFLIREMRTQPSLVQTOEQEYELVYNVAVLELFKROMVIRD 300
DB 240 vdytwmlldgiiipknfsvnlqemrtqpslvqtqeqyelvysavlelfkthmdvisd 299
QY 301 KHSGETSQAKHICPEKNHTLQADSYSNLPKSTTKAAKMMNQ---ORTKMEIKESSSDF 357
DB 300 nhlgreiaqcsipegsltveadscpldlpknamrdvkttnqhsqqaeeastggsslg 359
QY 358 RTSEISAKELVLHPAKSSTSFDFLELNYSFKNADTTMKWTKAPPIVIGEPQLKHQSLD 417
DB 360 rtstmaeeelvlhsaksspsfnclelncgcnknavitrngqaraspvvgelqkyqsl 419
QY 418 LGSLLFEGCSNKPVNAGRFNSKVPITRTKSTPELIIQORETKVDSENFYSLESQP 477
DB 420 fgmllfgscpsalpinatdryhnskpgvkrkatpfellqgrktndlavdgdfsclesql 479
QY 478 HDSCFV-EMQAOQVMHVSSAELNYSLPYDSKHQIRNASNVKHDSSALGVYSYIPLVENP 536
DB 480 hehysirelqvrvahvsseelnyslp-----gacdacvprhspgairvhllytslaedp 534
QY 537 YFSWPPPSGTSKMSLDLPEKQDGTVPFSSLLPTSTSLFSYVNSHSSLSLNSPTNSSL 596
DB 535 yfssppnsadsksmsfdlpekqdgatspgallpassttsffynphdslymntltsfpp 594
QY 597 LNOESAVLATAPRIDDEIPPPPLVPTPESFIVVEEAGEFSPNVPKSLSSAVKVIKIGTSLE 656
DB 595 lndetaveapsrtddeipplpertsfivveeagepsprvtesl--plvvtfgaspe 652
QY 657 WGGTSPKPFDDSVILRPKSVKLRSPKSGKN 688
DB 653 csqtsenks-hdsvgftpsknvklrspksdrh 683
RESULT      6
AAAM51205
ID  AAAM51205 standard; Protein; 780 AA.
XX
AC  AAAM51205;
XX
DT  19-DEC-2001 (first entry)
XX
DE  Human PTP-PEST Genbank Accession Number XP034191.
XX
KW  PTP; protein tyrosine phosphatase; tyrosine phosphorylated polypeptide;
KW  dephosphorylation; phosphotyrosine; human; PTP-PEST.
XX
OS  Homo sapiens.
XX
XX  WO200161031-A2.
XX
XX  23-AUG-2001.
XX
XX  13-FEB-2001; 2001WO-US05180.
XX
XX  14-FEB-2000; 2000US-0181769.
XX
XX  (CEPT-) CEPTYR INC.
PA

```

```
XX Flint AJ, Cool DE;
XX WPI; 2001-570570/64.
XX Screening assays to identify agents that alter protein tyrosine
XX phosphatase (PTP) binding to, and PTP-mediated catalytic
XX dephosphorylation of phosphotyrosine peptide substrates -
XX Example 1; Page -; 79pp; English.
XX The invention relates to identifying agents which alter the interaction
XX between a protein tyrosine phosphatase (PTP) and a tyrosine
XX phosphorylated polypeptide using fluorescence energy signals. The methods
XX are useful for performing screening assay to identify agents that alter
XX PTP binding to and PTP-mediated catalytic dephosphorylation of
XX phosphotyrosine peptide substrates. The present sequence is not given in
XX the specification but is that of human PTP-PEST protein sequence taken
XX from Genbank (Accession Number; XP034191). The present sequence was used
XX to generate mutants D199A (AAM51201) and C231S (AAM51202) as described in
XX the specification.
XX Note: An alternative sequence for human PTP-PEST is given in figure 1 of
XX the specification (AAG78281).
XX Sequence 780 AA;
SQ
Query Match 32.0%; Score 1158; DB 22; Length 780;
Best Local Similarity 37.2%; Pred. No. 1e-89;
Matches 291; Conservative 111; Mismatches 237; Indels 144; Gaps 25;
QY 1 MDQREIQLFLDEAOSKKIT-----KEEFANEFKLKROSTKYKADKTYPTTVAENAKNIK 56
DB 1 meqveillrficrvqamkspdhngednfardfmriristkyrteklyptatgekeenvk 60
QY 57 KRYRQDILPYDSRVLSLITSDDESSYINANFIVGVGPKAYIATQGLSTLLDFWRM 116
DB 61 krrykdilpfdhsrvkltiktqsgdsyinanfivgvpkayvatqplantvdfwrm 120
QY 117 IWEYSVLIIVMACMEYEMGKCEKRYWABPGEMQLEFGPFVSCAEAKRKSDYIIRTLKV 176
DB 121 iweynvliivmacrefemgkceerywlygedpdtfapfkiscdeqartdyfirtlli 180
QY 177 KFNSERTTIQYFHYKNWPDHVPSSIDPLILEIIMVRCYQEDSDVPICHCAGCGRTGV 236
DB 181 efqnesrilyqthvwnpvdhvpssfdslldmislmrkyqebedvpicihcsagcgrtga 240
QY 237 ICAIVDYTWMLLKDGIIPENFSVSLIREMPTQPSLQVQYELVYNNAVLELFRQMD 296
DB 241 icai-dytwnllkagkipeefnvnligemrtgrhsavqtkeqyelvhraiaqlfekqliq 299
QY 297 V-----IRD--KHSGETSQAKHCIPKNTLQADSYSPNLPKSTTKAAKMMNQQRTK 346
DB 300 lyeihgaqkiadgvneintemmvssiepek-----dspppkprrtr-----clvegda 350
QY 347 MEIKE-----SSSFDRFTSEISAKEEL-----VLHPAKSSTSFDFLELNY 387
DB 351 eellqppehpvppliltpspasfptvtvwdndryhpkvplhmvsseqhsadlnnys 410
QY 388 -----FDKNADTMKQYTAFFIVGEPKQKHOSLDLGSLLFEG----- 425
DB 411 kstelpgkneestieqdkklernalsfeikvplqegp-----ksfdgntllnrghaikiks 466
QY 426 ---C-----SNSKP-----VNAAGRYNSKVPIRTTKS-----TPFELIQRET 462
DB 467 aspciadkiskqogelsadinvdgtsgnscvdcsvtqskvsvtppesqnsdtpprpdl 526
QY 463 EVDSKENFYLEQPHSDCFVEMQAQKVMHVSFAELNY-----SLPYDSKHQIRNASW 516
DB 527 pldekghvtfwsfngpna-----ipidlsegnssdinyqtrktvsltpspqtqvdpdlv 582
QY 517 KHHDSALGVYSIPLVENPVPSSWPPSGTSSKMSLDLPEK-QDGTVPFSSLLPLPSTSL 575
DB 517 KHHDSALGVYSIPLVENPVPSSWPPSGTSSKMSLDLPEK-QDGTVPFSSLLPLPSTSL 575
```

```
DB 583 dhndts-----plfrtplsfnplhdsds-----dsdernsgavtqnktnistasat 630
QY 576 FSYNHSLSLNSTNINISLLNQESAVLATAPR-----IDDEIPPLPVRTPESFIVVEE 631
DB 631 vsaatstesistrkvlpmislarhniagtthsgaekdvdsedspplpertsfvlase 690
QY 632 AGEFSPNVPKSLSSAVKVKIGTSLFWG-----GTSEPKKFDSDSVILRPSKSVKLSRPSKG 686
DB 691 -----hntp-----vrsewselqserseqkkseg-----litsenekcdhpagg 730
QY 687 KNF 689
DB 731 lhy 733
RESULT 7
AAY67252
ID AAY67252 standard; protein; 780 AA.
XX AC AAY67252;
XX DT 05-APR-2000 (first entry)
XX DE Human protein tyrosine phosphatase (PTP) PEST amino acid sequence.
XX KW Protein tyrosine phosphatase; PTP; PEST; paxillin binding; adhesion;
XX KW cell migration; division; cytosolic; antiinflammatory; angiogenesis;
XX KW cancer; enzyme substrate identification; human.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Domain 333..340
FT /label= Pro_1
FT /note= "Proline rich domain 1"
FT Domain 356..365
FT /label= Pro_2
FT /note= "Proline rich domain 2"
FT Domain 519..528
FT /label= Pro_3
FT /note= "Proline rich domain 3"
FT Domain 674..680
FT /label= Pro_4
FT /note= "Proline rich domain 4"
FT Domain 769..776
FT /label= Pro_5
FT /note= "Proline rich domain 5"
XX WO9961467-A2.
XX PN 02-DEC-1999.
XX PD 21-MAY-1999; 99WO-CA00461.
XX PF 21-MAY-1998; 98CA-2238654.
XX PR 11-DEC-1998; 98US-0111993.
XX PA (UYMC-) UNIV MCGILL.
XX PI Tremblay ML, Cote J, Angers-Lousteau A, Charest A;
XX WPI; 2000-097104/08.
XX Novel therapeutic agents for treating diseases which are related to
XX cell proliferation, migration, inflammation and angiogenesis especially
XX cancer -
XX Claim 4; Fig 24; 91pp; English.
XX This is the amino acid sequence of human protein tyrosine phosphatase
XX (PTP) PEST. PTP-PEST is a soluble PTP that is ubiquitously expressed
XX throughout embryonic development and in murine adult tissues. The
XX N-terminal portion of the enzyme encodes for the catalytic domain, while
```



```

691  -----http-----vrsesqlqserseqkseg---ltsenekdhpagg 730
Db

```

DE	Human PTP-PEST mutant D199A.
XX	
KW	PTP; protein tyrosine phosphatase; tyrosine phosphorylated polypeptide
KW	dephosphorylation; phosphotyrosine; human; PTP-PEST; mutein; mutant.

XX
KW
OS
OS
OS

dephosphorylation; phosphotyrosine; human; FGF-FRS1; mutant.
Homo sapiens.
Synthetic.

AA	Key	Location/Qualifiers	
FT	Misc-difference	199	
FT		/note= "Wild-type Asp substituted by Ala"	
PN	WO200161031-A2.		
XX	23-AUG-2001.		
PD			
XX	13-FEB-2001; 2001WO-US05180.		
XX			
XX	14-FEB-2000; 2000US-0181769.		
PR			
XX	(CEPT-) CEPTYR INC.		
XX			
XX	Flint AJ, Cool DE;		
XX			
XX	WPI; 2001-570570/64.		
DR			
XX			
PT	Screening assays to identify agents that alter protein tyrosine		
PT	phosphatase (PTP) binding to, and PTP-mediated catalytic		
PT	dephosphorylation of phosphotyrosine peptide substrates -		
XX			
XX	Example 1; Page -; 79pp; English.		
PS			
XX			
CC	The invention relates to identifying agents which alter the interaction		
CC	between a protein tyrosine phosphatase (PTP) and a tyrosine		
CC	phosphorylated polypeptide using fluorescence energy signals. The met		
CC	are useful for performing screening assay to identify agents that alt		
CC	PTP binding to and PTP-mediated catalytic dephosphorylation of		
CC	phosphotyrosine peptide substrates. The present sequence is that of a		
CC	human PTP-PEST D199A mutant.		
CC	Note: The present sequence is not given in the specification but is		
CC	derived from the wildtype human PTP-PEST sequence (AAM51205) from		
CC	Genebank (Accession Number: XP034191). An alternative sequence for hu		
CC	PTP-PEST is given in figure 1 of the specification (AAG78281).		
XX			
XX	Sequence	780 AA;	
SQ			
Query Match			
Best Local Similarity 31.8%; Score 1150; DB 22; Length 780;			
Matches 290; Conservative 111; Mismatches 238; Indels 144; Gaps			
QY	1	MDQREIIQKFLDEAQAQSKKIT- ---KEEFANEFKLKRSQTKYKADKTYPTTVAENAKNIK 56	
Db	1	meqveilrkfiqrqamkspdhngednfdmfrlrlstkyrtekiyptatgeenvk 60	
QY	57	KNRYKDILPDYDYGVELSLTSDSDSYINANFIKGYGPKAYIATQGPLSTTLLDFWRM 116	

```
Db 61 knrykdilpfdhsrvkltlktspqdsdyinanfikvgypkayvatqgplantvidfwm 120
QY 117 IWEYSVLIIVMACMEYEMGKKKERYWAEPMQEMQLEFGPFSVSCBAERKSDYIIRTLKV 176
Db 121 Iweynvliivmacrefemgrkcerywplygedpitfapfkiscdeqartdyfirtlil 180
QY 177 KFNSETRTIYQHYKNWPDHVPSSIDPILILEIWDVRCYQEDDSVPICIHCSAGCGRTGV 236
Db 181 efqnesrrlyqfhyvnpahdvpssfdslmislmrkygehedvpicihcsagcgrtga 240
QY 237 ICAIVDYTWMLKDGILIPENFSVSLIREMRTQPSLQTOQOYELVYNAVLELFRQMD 296
Db 241 lcal-dytnlllkagkipeefnvnfligemrtqrsavqtkeqyelvhratqlfekqlq 299
QY 297 V-----IRD--KHSGTESQAKHCIPKNTHTLQADSYSPNLKSTTKAAKMMNQORTK 346
Db 300 Iyeihgaqkiadgvneintemvssiepek-----dspppkpptrs----clvegda 350
QY 347 MEIKE-----SSSDFPRTSEISAKEL-----VLHPAKSSTSFDFLELNY 387
Db 351 eellqppephvpplltppspafptvtvwdndryhpkpvlhmvsseqhsadlnrnys 410
QY 388 -----FDKNADTTMKWTKAFPIVGEPIQKHQSLDLSLLFEG-----425
Db 411 ktelpgknestieqldkklernlsfeikkvplqegp-----ksfdgntllnrghaik 466
QY 426 ---C---SNSKP-----VNAAGRYFNKSVPIRTTKS-----TPFELIQORETK 462
Db 467 aspiadkiskpqelssdinvgdtsgnscvdcsvtkgnkvsvtppesqnsdtprrdrl 526
QY 463 EYDSENFSYLSQPHDSCFVEMQAKVMHVSSEALNY-----SLPYDSKHQIRNASV 516
Db 527 pldkghvtwsthgpena---ipldiseqnsdinyqtrktvsltpspqtqvtpdlv 582
QY 517 KHDSSALGVSYIPLVNPENYFSSPPSGTSSKMSLDLPEK-ODGTVFPSSLLPTSSSL 575
Db 583 dhndts-----plfrclftnplhsdds-----dsdernsdgavtgnkntasat 630
QY 576 FSYNSHSLNSLNSPTNISLLNQESAVLATAPR-----IDDEIPPLPVRTPESTFV 631
Db 631 vsaatstesistrkvlpmislarhniagtthsgaekdvdsedspplpertsfvlase 690
QY 632 AGEFSNPVKLSAASVAVKIGTSLEWG-----GTSEPKKFDDSVILRPKSVKLSPKSG 686
Db 691 -----hntp-----vrsewselqserseqkkseg---litsenekdchpag 730
QY 687 KNF 689
Db 731 lhy 733

RESULT 9
AAM51202
ID AAM51202 standard; Protein; 780 AA.
XX
AC AAM51202;
XX
XX 19-DEC-2001 (first entry)
DE
DE Human PTP-PEST mutant C231S.
XX
KW PTP; protein tyrosine phosphatase; tyrosine phosphorylated polypeptide;
KW dephosphorylation; phosphotyrosine; human; PTP-PEST; mutain; mutant.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH
FT Misc-difference 231
XX
XX /note= "Wild-type Cys substituted by Ser"
XX
XX WO200161031-A2.
```

```
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US05180.
XX
PR 14-FEB-2000; 2000US-0181769.
XX
PA (CEPT-) CEPTYR INC.
XX
PI Flint AJ, Cool DE;
XX
XX WPI; 2001-570570/64.
XX
XX Screening assays to identify agents that alter protein tyrosine
XX phosphatase (PTP) binding to, and PTP-mediated catalytic
XX dephosphorylation of phosphotyrosine peptide substrates -
XX
XX Example 1; Page -; 79pp; English.
XX
XX The invention relates to identifying agents which alter the interaction
XX between a protein tyrosine phosphatase (PTP) and a tyrosine
XX phosphorylated polypeptide using fluorescence energy signals. The methods
XX are useful for performing screening assay to identify agents that alter
XX PTP binding to and PTP-mediated catalytic dephosphorylation of
XX phosphotyrosine peptide substrates. The present sequence is that of a
XX human PTP-PEST C231S mutant.
XX Note: The present sequence is not given in the specification but is
XX derived from the wildtype human PTP-PEST sequence (AAM51205) from
XX Genbank (Accession Number; XP034191). An alternative sequence for human
XX PTP-PEST is given in figure 1 of the specification (AAG78281).
XX
SQ Sequence 780 AA;
XX
XX Query Match 31.8%; Score 1148; DB 22; Length 780;
XX Best Local Similarity 37.0%; Pred. No. 7.4e-89;
XX Matches 290; Conservative 111; Mismatches 238; Indels 144; Gaps 25;
QY 1 MDQREILQKFLDEAQSCKIT----KEEFANEFKLKROSTKYKADKYPTTVAENAKNIK 56
Db 1 meqveillrkfqrvcamkspdhngednfardfmrllstkyrtekiyptatgekenvk 60
QY 57 KNRYKDILPYDSRVSELSLITSDSSYINANFIKGVGPKAYIATQGLSTLLDFWRM 116
Db 61 knrykdilpfdhsrvkltlktspqdsdyinanfikvgypkayvatqgplantvidfwm 120
QY 117 IWEYSVLIIVMACMEYEMGKKKERYWAEPMQEMQLEFGPFSVSCBAERKSDYIIRTLKV 176
Db 121 Iweynvliivmacrefemgrkcerywplygedpitfapfkiscdeqartdyfirtlil 180
QY 177 KFNSETRTIYQHYKNWPDHVPSSIDPILILEIWDVRCYQEDDSVPICIHCSAGCGRTGV 236
Db 181 efqnesrrlyqfhyvnpahdvpssfdslmislmrkygehedvpicihcsagcgrtga 240
QY 237 ICAIVDYTWMLKDGILIPENFSVSLIREMRTQPSLQTOQOYELVYNAVLELFRQMD 296
Db 241 lcal-dytnlllkagkipeefnvnfligemrtqrsavqtkeqyelvhratqlfekqlq 299
QY 297 V-----IRD--KHSGTESQAKHCIPKNTHTLQADSYSPNLKSTTKAAKMMNQORTK 346
Db 300 Iyeihgaqkiadgvneintemvssiepek-----dspppkpptrs----clvegda 350
QY 347 MEIKE-----SSSDFPRTSEISAKEL-----VLHPAKSSTSFDFLELNY 387
Db 351 eellqppephvpplltppspafptvtvwdndryhpkpvlhmvsseqhsadlnrnys 410
QY 388 -----FDKNADTTMKWTKAFPIVGEPIQKHQSLDLSLLFEG-----425
Db 411 ktelpgknestieqldkklernlsfeikkvplqegp-----ksfdgntllnrghaik 466
QY 426 ---C---SNSKP-----VNAAGRYFNKSVPIRTTKS-----TPFELIQORETK 462
Db 467 aspiadkiskpqelssdinvgdtsgnscvdcsvtkgnkvsvtppesqnsdtprrdrl 526
```

Qy	463	EVDSENFSYLESQPHDSCFEVQAOKVMHVSSAELNY-----SLTPYDSKHQIIRNASNV	516
Db	527	pldgkhtvtfshgpena----ipldlsegnsndinyqtrktvsftpttqqvetpdlv	582
Qy	517	KHHDSSALGVYSIPILVENPYFSSWPSPGTSKMSJDLPEK-QDGIVFPSSLPTPSTSYSL	575
Db	583	dhnts-----plfrtplstfnlhadds-----dsdermsdgaavgtnkstasat	630
Qy	576	FSYNHGSHLSLNSPNISLLNQESAVIATAPR---IDDEIPPLPVRTPESTIVEE	631
Db	631	vsaatesistrlkvlpmsiarhnlagtthsgaekdvdsedpplpertpestivfase	690
Qy	632	AGEFPSNVPKSLSSAVKVKGITSLEWG-----GTSEPKKFDSDVILLRPSEKSVKLRSKSG	686
Db	691	-----hntp-----vrsewselqsqrseqkkseg---litsekcdhpagg	730
Qy	687	KNF	689
Db	731	iHy	733

RESULT	10
--------	----

AAY67253
ID AAY67253 standard; protein; 780 AA.

interfering with the binding of PTP-PEST to signalling molecules that are involved in cell migration, adhesion or division. The compound can be derived from minimal sequences found in binding sites of PTP-PEST. The invention also relates to a method for finding a genuine substrate for an enzyme in a cell that normally expresses the wild type form of the enzyme. A mutant version of PTP-PEST (represented by this sequence) is used in this method. The compounds have cytostatic and antiinflammatory activity. The compounds are used for making medicaments for treating a disease related with cell proliferation, migration, cell inflammation and angiogenesis, especially cancer. The novel method is used for identifying a genuine substrate for an enzyme.

Note: The present sequence is not shown in the specification but it has been derived from the PTP-PEST sequence (see AAY67252), shown in figure 24.

```

Query Match      31.7%; Score 1147; DB 21; Length 780;
Best Local Similarity 37.0%; Pred. No. 9e-89;
Matches 290; Conservative 111; Mismatches 238; Indels 144; Gaps 25;

QY      1 MDQREILQKFLDEAQSKIT----KEEFANFELKLKROSTRKYKADKYPTTVAENAKNIK 56
DB      1 meqveilrkfiqrvaqmsphngednfardfmrlrlstkyrtekytptatgekenvk 60
QY      57 KNYRKDILPYDYSRVEISLITSDSDSSYINANFIKVGYPKRAYITGLPTLTLDFWRM 116
DB      61 knrykdilpfdhsrvtkltlpsqdsdyinanfikvygpkayvatgplantvldfrm 120
QY      117 IWEYSVLIIWACMEYEMGKKCCERYWAEPCEMQLEPGFPFVSCEAEKRRKSDYIIRTLKV 176
DB      121 iweynvliiwmacreferemgrkcceryplygedpittfapfkiscedeqardtyfirtlll 180
QY      177 KFNSERTIIQFYHKNPDPDHPSSIDPILELIIMWRCYQBDSDVPICIHCSAGCGRTGV 236
DB      181 efqnesrlrqfyhynwpdhvpsfdisldmismrkygehedvpicihesagcgrtga 240
QY      237 ICALVDYTWMLLKDGIIIPENPSVFSLSIREMTPQPSLVOEQVELYVNAVLEIFPKROMD 296
DB      241 ical-dytwllkagkipeefnvnlligemrtqhsavqtkegyvelvhrailaqfekqlq 299
QY      297 V-----LRD--KHSGTESQAKHCPIEKNHTLQADQSYSPNLKSTTAKAKMNOQRTK 346
DB      300 lyeihgaqkiadgveintemmissiepeq----dsppkpprtss----clivegak 350
QY      347 MEIKE-----SSSDFRTISEAKEEL-----VLHPAKSSTSFDFLELNY 387
DB      351 eelilqppehpvpilltpspsaftvtvqgdndryhpkpvlhmvsseqhsadlnrny 410
QY      388 -----FDKNADTTMKWQTKAPPIVGEPLQKHQSLDGLSLLFEG----- 425
DB      411 kstelpgknestieqldkklernisfeikkyplqegp----ksfdgntilnrghaikls 466
QY      426 --C---SNSKP-----VNAAGRYFNFSKVPITRTKS-----TPFELTQQRETK----- 462
DB      467 aspciaakickpqlssdlngvdtisqnsdcsvtqsnkvsvtpeesqnsdtpprpdl 526
QY      463 EVDSKENFSYLESOPHSCFCFVMOAQKVMHVSSAELNY-----SLPYDSKHQIRNASNV 516
DB      527 pldekghvtsfghpna----ipidlsegnssdinyqtrkvtksaltsptpsptqvetpdlv 582
QY      517 KHHDSALGVYSYIPLVBNYPFSSWPSPGTSKKMSLDLPEK-ODQTVFPSSLLPTSSLSL 575
DB      583 dhndts-----plfrtplisfncplhdsds-----dsdernsdgavcqnktnlstasat 630
QY      576 FSYYNHSSLSLNSPTNISLLNOCESAVLATAPR-----IDDEIPEPLPVPRPEPSFIVVEE 631
DB      631 vsaatstesistrkvlpmnsiarhniagtthsgackdvdsdpsppilpertpesfvase 690
QY      632 AGEFSPNVPKSLSGAVKVKIGTSLIEWG-----GTSEPKKFDDSVILTRPSKVULRSPKSG 686
DB      691 -----http-----vreswelsqserceqkkseq-----litseneckdhpagg 730

```

QY 687 KNF 689
 Db 731 lby 733

RESULT 11
 AAY67250
 ID AAY67250 standard; protein; 775 AA.
 AC AAY67250;
 XX
 XX 05-APR-2000 (first entry)
 DT
 DE Mouse protein tyrosine phosphatase (PTP) PEST amino acid sequence.
 XX
 XX Protein tyrosine phosphatase; PTP; PEST; paxillin binding; adhesion;
 KW cell migration; division; cytosolic; antiinflammatory; angiogenesis;
 KW cancer; enzyme substrate identification; mouse.
 XX
 OS Mus sp.
 FH
 FH Key Location/Qualifiers
 FT Domain 332..339
 FT /label= Pro_1
 FT /note= "Proline rich domain 1"
 FT Domain 355..364
 FT /label= Pro_2
 FT /note= "Proline rich domain 2"
 FT Domain 519..528
 FT /label= Pro_3
 FT /note= "Proline rich domain 3"
 FT Domain 675..681
 FT /label= Pro_4
 FT /note= "Proline rich domain 4"
 FT Domain 764..771
 FT /label= Pro_5
 FT /note= "Proline rich domain 5"
 XX
 XX WO9961467-A2.
 PN
 PN 02-DEC-1999.
 PD
 XX 21-MAY-1999; 99WO-CA00461.
 XX
 XX 21-MAY-1998; 98CA-2238654.
 PR
 PR 11-DEC-1998; 98US-0111993.
 XX
 XX (UYMC-) UNIV MCGILL.
 PA
 PI Tremblay ML, Cote J, Angers-Lousteau A, Charest A;
 PI WPI; 2000-097104/08.
 XX
 XX Novel therapeutic agents for treating diseases which are related to
 PT cell proliferation, migration, inflammation and angiogenesis especially
 PT cancer -
 XX
 XX Claim 4; Fig 24; 91pp; English.
 PS
 XX This is the amino acid sequence of mouse protein tyrosine phosphatase
 CC (PTP) PEST. PTP-PEST is a soluble PTP that is ubiquitously expressed
 CC throughout embryonic development and in murine adult tissues. The
 CC N-terminal portion of the enzyme encodes for the catalytic domain, while
 CC the C-terminal portion is composed of 5 proline rich domains, and a
 CC binding site for the adaptor protein Shc. The pro 2 domain is required
 CC for paxillin binding, and the synthesis of mutant PTP-PEST have shown
 CC that proline 362 is important for paxillin binding activity. The
 CC invention relates to a compound that is capable of interfering with the
 CC binding of PTP-PEST to signalling molecules that are involved in cell
 CC migration, adhesion or division. The compound can be derived from
 CC minimal sequences found in binding sites of PTP-PEST. The invention also
 CC relates to a method for finding a genuine substrate for an enzyme in a

CC cell that normally expresses the wild type form of the enzyme. A mutant
 CC version of PTP-PEST (see AAY67251 and AAY67253) is used in this method.
 CC The compounds have cytostatic and antiinflammatory activity. The
 CC compounds are used for making medicaments for treating a disease related
 CC with cell proliferation, migration, inflammation and angiogenesis,
 CC especially cancer. The novel method is used for identifying a genuine
 CC substrate for an enzyme.
 XX
 SQ Sequence 775 AA;

Query Match 31.2%; Score 1127.5; DB 21; Length 775;
 Best Local Similarity 36.6%; Pred. NO. 4.2e-87;
 Matches 280; Conservative 109; Mismatches 166; Indels 209; Gaps 28;

QY 1 MDQRETLQKFLDEAQSCKIT---KEEFANEFLKLRQSTKYKADKTYPTTVAENAKNIK 56
 Db 1 meqvellrrlfrgvdamkspdhngednrfmrirrlstkyrtklyptatgekeenvk 60
 QY 57 KNRKYDILPYDYSRVLSLITSDSSYINANFIKVGYPKAYIATQGPLSTLLDFWRM 116
 Db 61 knrykdilpfdhsrvkltlktpsqdsdynanfikvgypkayvatqgpfntvidfwm 120
 QY 117 IWESVLIIVMACMEYEMGKKCERYWAPGEMQLEFGPFSVSCAEKRSKIYIIRLKV 176
 Db 121 iweynvviivmacrefemgrkkcerywplygedpitfapfkiscenegartdyfirtll 180
 QY 177 KENSETRIYQPHYKNWPDHVPSSIDPILWVRCYQEDDSVPICIHCSAGGCGTV 236
 Db 181 efqnesrrlyqfhywnpdhvpssfdslmslmrkyqehedvpichcsagcgtrga 240
 QY 237 ICAIVDYTMMLKDGIIIPENFVSFLIREMRTORFSLVQTOBYELVYNAVLELFKQMD 296
 Db 241 icai-dywnllkagkipeeovfnliqemrtqrhsavqtkeqyelvhrailfekqlq 299
 QY 297 V-----IRDKHSGTESQAKHCIPKNTLOADSYSPNLPKSTTKAAKMNQQRKME 348
 Db 300 lyeihgaqkiadgneittgtvmssidse---kqdsppkpprttrs---clvegdakee 351
 QY 349 IKE-----SSSFDFRTS-----EISAKEELVLHPAKSSTSF 379
 Db 352 ilqppephvppiltpspsaftvtvwdsdryhpkpvlhmaspeq---hpa----- 402
 QY 380 DFLELNYSDKNADTTMKW-----QTKAFPIVGEPLQK-HQSLDLGSLLF 423
 Db 403 ---dlrnsydkasad---qwgksesaiehidkklernlsfeilkvplqeqpkfsgdntlln 456
 QY 424 EGCNSKPVNNAAGRYFNKVPITRTKSTFEL-----IQORETKEV----- 464
 Db 457 rg--haikiksa-----sssvvdrtskqelsagalkvddvsqncadcsaahsraae 508
 QY 465 DSKENFSYLESOPHDSCFVEMQAQKVMHVSAAELNYSL-----PYDSK 507
 Db 509 sseeesnshtpprdc---lpldkghvt-----wslhgenatpvpdpdgkspdnhs 560
 QY 508 HOIRNASN-----VKHDSALGVYSYIPLVENP-----YFSWPPSGTGS 547
 Db 561 qtlktvstspnstaeaeahdltehnss-----pllkapslfcnpilhsddwhsdggs 612
 QY 548 SKMSLDLPKQDGTVPFSSLLPTSTSLFSYVNSHSSLSLNSPTNISLLNQESAV---- 603
 Db 613 s-----dgav-----trnktsistat-vspassaesaachrrv 645
 QY 604 -----LATAPR-----IDDEIPPLPVRTPESFIVE 630
 Db 646 lpmstarqevagtphsgaekdadvseesppplpertsfviad 689

RESULT 12
 AAY67251
 ID AAY67251 standard; protein; 775 AA.
 XX
 AC AAY67251;


```
XX 08-MAR-2000; 2000WO-US05881.
XX 12-MAR-1999; 99US-0124270.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM;
XX WPI; 2000-611515/58.
XX N-PSDB; AAF21801.
XX New human breast and ovarian cancer associated gene sequences and the
XX polypeptides encoded by these genes, useful in the prevention,
XX treatment and diagnosis of cancer, immune disorders, cardiovascular
XX disorders and neurological diseases -
XX Claim 11; Page 1045-1047; 1299pp; English.
XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
XX proteins AAB58711 - AAB59128. The DNA and protein sequences are
XX associated with breast and ovarian cancer. Included in the invention are
XX sequences AAF22032 - AAF22040 and AAB59129 which are used in the
XX isolation and characterisation of the DNA and protein sequences of the
XX invention. The breast and ovarian cancer associated DNA, protein, agonist
XX or antagonist sequences exhibit cytostatic; immunosuppressive;
XX neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
XX antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
XX antibacterial; antifungal; antiparasitic and cardiant activity. The
XX polynucleotide and protein sequences are used in the diagnosis of cancer,
XX particularly breast and ovarian cancer. The nucleic acid sequences,
XX proteins, agonists and antagonists may also be used in the diagnosis,
XX prevention and treatment of immune disorders e.g. Addison's disease,
XX allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
XX diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
XX arthritis and ulcerative colitis; cardiovascular disorders such as
XX myocardial ischaemias; wound healing; neurological diseases such as
XX cerebral anoxia and epilepsy; and infectious diseases.
XX Sequence 331 AA;
XX
XX Query Match 29.3%; Score 1058; DB 21; Length 331;
XX Best Local Similarity 99.0%; Pred. No. 8.9e-82;
XX Matches 206; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 478 HDSCFVMOAQKVMHVSSEALNYSPLPDSKHOIRNASNVKHHDSALGVSXIPLVENPY 537
XX Db 1 hdsfcvmaqakvmhvssealnysplpdskhqirnasnvkhhdsalgvsiplvenpy 60
XX
XX QY 538 FSWPPSGTSSKMSLDLPKQDGTVPSSLLPTSTSLFSYNSHSSLSLNSPTNISLL 597
XX Db 61 fswppsgtsskmsldlpkqdgvtvpssllptstslfsyynshdslslnsptnissll 120
XX
XX QY 598 NQESAVLATAPRIDDDIPPLVPRTPESTIVVEEAGEFSPNPKSLSSAVKVKIGTSLBW 657
XX Db 121 nqesavlatapridddipplvprtpestivveeagefspnypklsissavkvkigtlslew 180
XX
XX QY 658 GGTSEPKKEDDVSILRPKSKVLRSKPS 685
XX Db 181 ggtsepkkddsvilrpsksvklrpsks 208
XX
XX RESULT 14
XX AAG78281
XX ID AAG78281 standard; Protein; 312 AA.
XX AC AAG78281;
XX
XX 19-DEC-2001 (first entry)
XX DT Human PTP-PEST.
XX
```

```
XX PTP; protein tyrosine phosphatase; tyrosine phosphorylated polypeptide;
XX dephosphorylation; phosphotyrosine; human; PTP1B; mouse; fruit fly;
XX yeast.
XX OS Homo sapiens.
XX PN WO200161031-A2.
XX PD 23-AUG-2001.
XX PF 13-FEB-2001; 2001WO-US05180.
XX PR 14-FEB-2000; 2000US-0181769.
XX PA (CEPT-) CEPTYR INC.
XX PI Flint AJ, Cool DE;
XX WPI; 2001-570570/64.
XX Screening assays to identify agents that alter protein tyrosine
XX phosphatase (PTP) binding to, and PTP-mediated catalytic
XX dephosphorylation of phosphotyrosine peptide substrates -
XX Disclosure; Fig 1; 79pp; English.
XX The invention relates to identifying agents which alter the interaction
XX between a protein tyrosine phosphatase (PTP) and a tyrosine
XX phosphorylated polypeptide using fluorescence energy signals. The methods
XX are useful for performing screening assay to identify agents that alter
XX PTP binding to and PTP-mediated catalytic dephosphorylation of
XX phosphotyrosine peptide substrates. The present sequence is that of a
XX catalytic domain of a PTP for comparison with human PTP1B (AAG78262).
XX SQ Sequence 312 AA;
XX
XX Query Match 25.8%; Score 933; DB 22; Length 312;
XX Best Local Similarity 54.2%; Pred. No. 3.9e-71;
XX Matches 182; Conservative 47; Mismatches 57; Indels 50; Gaps 6;
XX
XX QY 44 YPTTVAENAKNKKRYKDILPYDSRVLSLTSDESSYINANFKGVGPKAYIATQ 103
XX Db 2 yptatgekeenvkknrykdilpfdhsrvkltktpsqdsdyinankfgyvgpkayvatq 61
XX
XX QY 104 GPLSTTLDFWRMIWEYSVLIIVMACMEYEMGKKCKERYWAEPEGMLFEGPFSVCEAE 163
XX Db 62 gplantvldfwrmiweysvliivmacmevemgkckckerywplygedpfitfapfkiscde 121
XX
XX QY 164 KRKSDYIIIRTLKVKFNSETRTIYQFHYKNWPDHDPVSSIDPILELIWDRVCYQEDDSVPI 223
XX Db 122 qartdyfi-lllelqnesrrlyqfhyvnpdhdvpsfidsldmslrmrkyqehedvpi 179
XX
XX QY 224 CLHCSAGCGRTGVCIAIVDYTWMLKDGIIIPENFVSFLIREMRTORPSSLVQTQOEYELV 283
XX Db 180 clhcsagcgrtgaicai-dytwnllkagkipeefnfnliqemrtgrhsavgtkeqyelv 238
XX
XX QY 284 YNAVLEFRQMDVIRDKSGTESQAKHCIPKKNHTLQADSYSPNLPKSTTKAAK----- 338
XX Db 239 hraiaqlfekdlq-----lyaidytwnl-----lkagkipeef 271
XX
XX QY 339 ----MMNQORTKMEIKESSSDFRTSEISAKBELVL 370
XX Db 272 nvfnliqemrtg-----rhsvavtkeqyel 296
XX
XX RESULT 15
XX AAB59383
XX ID AAB59383 standard; Protein; 253 AA.
XX XX
XX AC AAB59383;
XX
XX DT 21-MAR-2001 (first entry)
XX
```

XX Human protein tyrosine phosphatase #14.
DE Protein tyrosine phosphatase; human; mouse; fruit fly; PTP;
XX substrate trapping.
KW Homo sapiens.
OS Homo sapiens.
XX WO200075339-A1.
PN 14-DEC-2000.
XX 24-MAY-2000; 2000WO-US14211.
XX 03-JUN-1999; 99US-0137319.
PR 16-JUN-1999; 99US-0334575.
XX (COLD-) COLD SPRING HARBOR LAB.
PA Tonks NK, Zhang S;
PI WPI; 2001-080598/09.
XX New substrate trapping mutant protein tyrosine phosphatases (PTP) in
PT which the wild type PTP catalytic domain invariant aspartate is
PT replaced with an unphosphorylated amino acid, useful in gene therapy
XX Disclosure; Fig 1; 109pp; English.
XX The present invention provides substrate trapping mutant protein tyrosine
CC phosphatases (PTPs). They can be used to reduce the activity of tyrosine
CC phosphorylated proteins and to screen for modulators capable of altering
CC the binding of protein tyrosine phosphatases to their substrate. These
CC may be used in disease diagnosis and treatment.
XX SQ Sequence 253 AA;

Query Match 25.5%; Score 921.5; DB 22; Length 253;
Best Local Similarity 65.4%; Pred. No. 2.7e-70;
Matches 166; Conservative 41; Mismatches 44; Indels 3; Gaps 2;

QY 44 YPTTVAENAKNKKRYKDIILPYDYSVELSLITSDSSVINANFIKGVGPKAYIATQ 103
DB 2 YPTATGEKEENVKKNRYKDIILPDHRSVKITLTPSQDSQSYINANFIKGVGPKAYVATG 61

QY 104 GPLSTTLLDFWRMIWEYSVLIIVMACMEYEMGKKKCYRYNAEPCGEMOLEFGPFVSVCSEAE 163
DB 62 GPLANTVIDFWRMWEYNVNVIIVMACREFENGRRKKERYPLYGEDPITFAPFKISCDE 121

QY 164 KRKSDYIIRTLKVKFNSETRIYQFHYKKNPDPDHPVPSIDPILILEIWDVRCYQEDDSVPI 223
DB 122 qartdyfi--lllefqnesrrlyfhyvnpdhdvpsfsdmslmrkyqehedvpi 179

QY 224 CIHCSAGCGTGVICAVDVTWLLKGLIPENFVSFLIREMETORPSLVOTGOEYELV 283
DB 180 cihsagcgtrtgai-cal-dytwnllkagkipeefnvlqemrtqrhsavqtkeqyelv 238

QY 284 YNAVLELFKQMDV 297
DB 239 hraiaqlfekqlqi 252

THIS PAGE BLANK (USPTO)